May 2015

Probability Models for Wind-Penetrated Power Systems

Amir Hossein Shahirinia
University of Wisconsin-Milwaukee

PROBABILITY MODELS FOR WIND-PENETRATED POWER SYSTEMS

by

Amir Hossein Shahirinia

A Thesis Submitted in

Partial Fulfillment of the Requirements

for the Degree of

Doctor of Philosophy

In Engineering

at

The University of Wisconsin-Milwaukee

May 2015

ABSTRACT

PROBABILITY MODELS FOR WIND-PENETRATED POWER SYSTEMS

by

Amir Hossein Shahirinia

The University of Wisconsin-Milwaukee, 2015
Under the Supervision of Professor David Yu and Professor Ehsan Soofi

A major challenge with the increase in wind power generation is the uncertain nature of wind speed. So far the uncertainty about wind speed has been presented through probability distributions. However, the uncertainty about these wind speed models has not yet been considered. In this dissertation we use the Bayesian approach to taking into account the uncertainty inherent in the wind speed model. Also the existing models that consider the uncertainty of the wind speed primarily view the distributions of the wind speed over a wind farm as being homogeneous. The Bayesian predictive model of the wind speed aggregates the non-homogeneous distributions into a single continuous distribution. Therefore, the result is able to capture the variation among the probability distributions of the wind speeds at the turbines' locations in a wind farm. More specifically, instead of using a wind speed distribution whose parameters are known or estimated, the parameters are considered as random whose variations are according to probability distributions. In order to present the applications of developed uncertain models, we apply both non-Bayesian and Bayesian models to a well-known power systems problem known as Stochastic Economic Dispatching (S-ED). Traditionally, S-ED algorithms incorporate wind speed using a single point from a wind speed distribution to generate the resultant wind power as the input for the ED algorithm to produce the optimal combination of fossil fuel power generation. In this dissertation, we develop a new Stochastic Economic Dispatch algorithm, referred to as SEconD, for capturing the uncertainty induced by

the wind speed of the planning target time to economic dispatching output variables. SEconD uses the entire wind speed distribution as the input to generate the resultant wind power distribution rather than just a single point and produces data for estimating the probability distributions of optimal fossil fuel generation outputs, transmission loss, and total cost of power generation. Having distributions of optimal outputs enables a system operator to perform useful statistical analyses of the outputs.

To

my parents, my brother

and

my sister

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGMENTS

I would like to express my special appreciation and thanks to my advisors: Professor David Yu and Professor Ehsan Soofi. They have patiently mentored me throughout this project. I would like to thank them for all their encouragement and for guiding me to this point. Their advice has been priceless. I would also like to thank my other committee members: Professor Adel Nasiri, Professor Chiu Tai Law and Professor Junhong Chen.

A special thanks to my family. Words cannot express how grateful I am to my mother, father, sister, and brother for all of the sacrifices that they have made on my behalf. Their prayers for me have always given me energy to overcome challenges.

I would also like to thank my friends, especially Mark Anderson, Dr. Mojtaba Heydar and Dr. Ranga Tallam who encouraged and assisted me in reaching my goals.

I am grateful to **God** for bringing me through all the difficulties, for the day-to-day guidance, and for the opportunity to complete my degree program.

1. Chapter 1

Introduction

Environmental concerns have made wind power an appealing source of clean and renewable energy, and as this field continues to grow, calibrated and smart probabilistic forecasts can help to make wind power a more financially competitive alternative. However, this is challenging due to the uncertain nature of wind speed. Statistical methods have been applied to three different time scales: short-term, medium-term and long-term. Short-term wind speed forecasts play a central role in estimating various engineering parameters, such as power outputs, extreme wind loads, and fatigue loads [1]. The wind speed forecasts for this time scale are only a few hours ahead of target time [2]-[6]. Medium-term wind speed forecasts look several days ahead are generally based on weather prediction models, which can then be statistically post-processed [7]-[10]. Long-term wind speed forecasts require analysis of wind speed data over a number of years [11]. Probability distributions are used primarily to take into account the uncertainly of wind speed in all three time scales.

The quality of wind speed modeling depends on the suitability of the chosen probability models to describe the wind speed frequency distribution. An overview of the wind speed models used in recent literature is as follows. [12] and [13] used a Weibull distribution to forecast the wind speed and assessed wind energy potential. [14] compared fit of a Rayleigh distribution and another Weibull distribution to wind speed data and showed that the Weibull model provided a better fit. In [15], a wind speed distribution was shown to be satisfactorily described by a Lognormal distribution. In [16] Weibull and Lognormal distributions were used to fit wind speed frequencies

and concluded that the Weibull distribution better fit the data. [17] used Rayleigh, Weibull, and Gamma distributions to model wind speeds both on and offshore.

In addition Gumbel and the Generalized Extreme Value distributions were used to model extreme wind speeds [18]-[21].

The most commonly used models to describe wind speed distribution are Weibull and Gamma, both of which belong to the Generalized Gamma (GG) family. The GG family is flexible in that it includes several well-known models as subfamilies. The probability distributions of the GG are:

$$f(w\mid\alpha,k,c)=\frac{k}{c^{\alpha k}\Gamma(\alpha)}w^{\alpha k-1}e^{-\left(\frac{w}{c}\right)^{k}},w\geq 0,\alpha,k,c\succ 0, \tag{1-1}$$

and

$$F(w\mid\alpha,k,c)=\frac{\gamma\left(\alpha,\left(\frac{w}{c}\right)^{k}\right)}{\Gamma(\alpha)}. \tag{1-2}$$

The GG with $\alpha=1$ gives a Weibull with a shape parameter k and a scale parameter c. For k = 1, GG represents Gamma with a shape parameter α. The Lognormal distribution is also obtained as a limiting distribution when $\alpha\to\infty$. By letting $k=2$ it obtains a subfamily of GG which is known as the generalized normal distribution, GN(2α,c). The GN is itself a flexible family and includes Half-normal $\alpha=\frac{1}{2}$, Rayleigh $\alpha=1$, Maxwell–Boltzmann $\alpha=\frac{3}{2}$, and Chi $\alpha=\frac{k}{2},k=1,2,\ldots$. GG with $\alpha=k=1$ gives an exponential distribution [22], [23], [24] and [25].

The use of abovementioned models view the distributions of the wind speed over a wind farm as being homogenous. However, a wind farm has multiple turbines installed in different locations, each of which may have its own distribution model. For such cases a mixture model constructed by convex linear combinations of two or more distributions can be useful. The components of a

mixture model may belong to the same family of models, but do not necessarily have the same parameters. In the recent past, mixture distributions have been used to model wind speed characteristics to capture the variability of the wind speed in various seasons [26]-[28].

The PDF of a mixture model is given by:

$$f_{mix}(w) = \sum_{j=1}^{J} p_j f_j(w), w \geq 0, \quad (1\text{-}3)$$

where J is the number of models, $p_j \geq 0$ are the mixing parameter and $\sum_{j=1}^{J} p_j = 1$.

It should be noted that $f_j(w)$ can be either from different families of distributions or the same family but having different parameters values $\theta_j, f_j(w) = f(w|\theta_j)$.

A mixture model that combines distributions from the same family is expressed as follows:

$$f_{mix}(w) = \sum_{j=1}^{J} p(\theta_j) f_j(w|\theta_j), w \geq 0, \quad (1\text{-}4)$$

From the above model it can be interpreted that, for a randomly drawn wind speed data point, the probabilities of that wind speed data point belonging to model f_j is p_j.

The most basic wind speed mixture model is a bimodal model, which combines only two of the distributions, both of which are typically from the GG family [11], [22],[26]-[29]. For example a bimodal mixture PDF, $J = 2$, that combines two Weibull distributions $f(w|\theta_j)$ where $\theta_j = \{k_j, c_j\}$ with the probability of p_j is given by:

$$f_{mix}\left(w \mid k_1, c_1, k_2, c_2, p_1, p_2\right) = p_1 f\left(w \mid k_1, c_1\right) + p_2 f\left(w \mid k_2, c_2\right)$$
$$= p_1 \left\{ \left(\frac{k_1}{c_1}\right)\left(\frac{w}{c_1}\right)^{k_1 - 1} e^{-\left(\frac{w}{c_1}\right)^{k_1}} \right\} + p_2 \left\{ \left(\frac{k_2}{c_2}\right)\left(\frac{w}{c_2}\right)^{k_2 - 1} e^{-\left(\frac{w}{c_2}\right)^{k_2}} \right. \quad (1\text{-}5)$$

Similarly the CDF of the mixture model is as follows:

$$F_{mix}\left(w \mid k_1, c_1, k_2, c_2, p_1, p_2\right) = p_1 F\left(w \mid k_1, c_1\right) + p_2 F\left(w \mid k_2, c_2\right)$$
$$= p_1 \left\{ 1 - e^{-\left(\frac{w}{c_1}\right)^{k_1}} \right\} + p_2 \left\{ 1 - e^{-\left(\frac{w}{c_2}\right)^{k_2}} \right\}. \quad (1\text{-}6)$$

From the above model it can be interpreted that, for a randomly drawn wind speed data point, the probabilities of that wind speed data point belonging to the first Weibull model, $f_1(w|c_1,k_1)$, is p_1, and that of second Weibull model, $f_2(w|c_2,k_2)$ is p_2. It is obvious that the model captures the variation among probability distributions of the wind speeds at two different turbines' locations in a wind farm. This use of bimodal models is common in the literature [11], [22], [26]-[29].

Also θ_j can be viewed as random that varies according to probability distributions. In the Bayesian approach $p(\theta_j)$ is prior and f_{mix} is predictive model. The Bayesian predictive wind speed distribution is expressed as follows:

$$f_{mix}(w) = \int p\left(\theta\right) f\left(w \mid \theta\right) d\theta, \quad (1\text{-}7)$$

This model can aggregate a large number of non-homogeneous distributions into a single continuous distribution. In this case, a model which is a combination of a large number of models takes the form of a single mode density function due to the resultant close proximity of the modes one to another. The Bayesian predictive models have been widely used in the literature but not in

the context of Stochastic Economic Dispatching (S-ED). In this dissertation we will apply wind speed predictive models to the S-ED problem incorporating wind speed.

In electric power system literature, an Economic Dispatch (ED) model is used to minimize the cost of the production of power to be generated by several sources for a target time. With the increasing penetration of wind energy in power systems, two challenges facing system operators are how to incorporate the wind energy potential and how to reliably dispatch an optimal combination of all the available energy sources [30]-[37]. The primary problem associated with incorporating wind power into the ED model is the uncertainty about the wind speed for the dispatching target time, which is usually a day ahead. The stochastic aspect of wind speed in the ED context has been addressed in recent literature through the inclusion of the probability distribution function of wind power generation and the expected cost of over and/or under estimation of the wind speed into the ED model [32] and [34]. In addition, [35] has used Monte Carlo simulation and a forecasting model to generate short term wind speed forecasts for the ED model. These stochastic ED methods provide optimal combinations of the available energy sources to be dispatched. Typically they use a single randomly drawn data point from wind speed distribution as the input for the ED model and produce the optimal combination of outputs. However, producing PDFs for the optimal solutions in the context of the stochastic ED models has not yet been considered. This dissertation fulfills this void by introducing an algorithm that uses the entire wind speed distribution as the input rather than just a single point in order to generate the distributions of optimal outputs.

Producing the PDFs of the optimal solutions in the closely related problem of *optimal power flow* (OPF) is commonplace. The Monte Carlo simulation and some approximation methods are used for finding the solutions' PDFs in the OPF problem [38]-[42]. This dissertation proposes a Monte

Carlo (MC) simulation algorithm for producing PDFs of the optimal solutions of the stochastic ED models, hereafter referred to as SEconD.

This dissertation will describe the proposed algorithm, illustrate it through examples using non-Bayesian and Bayesian wind speed models as inputs, and discuss how a system operator can perform many useful analyses using the resultant outputs of SEconD. For example the system operator can do the economic dispatch of the fossil fuel power plants based on a range of wind speed with a given probability and determine the corresponding ranges of the optimal outputs of the fossil fuel power plants, minimum operation costs, and transmission losses. Similarly, for a sub-range of the optimal operation range of any fossil fuel power plant, the system operator will be able to determine its probability as well as the corresponding sub-ranges of the optimal operation ranges of the remaining generators and their probabilities. In addition, the system operator can perform "what if" analyses in order to compare the probability distributions of the optimal outputs of the fossil fuel power plants, the transmission losses, and optimal operation costs under various scenarios.

This dissertation is organized as follows. Chapter 2 describes the proposed SEconD algorithm. The Bayesian approach to modeling wind speed is presented in chapter 3. The SEconD algorithm will be applied to both a standard wind speed distribution, known as a non-Bayesian, and a Bayesian predictive wind speed models. Also a comprehensive comparison between non-Bayesian and Bayesian predictive wind speed models and their resultant outputs from the SEconD algorithm will be presented in chapter 4. Chapter 5 gives the concluding remarks along with future work.

2. Chapter 2

Stochastic Economic Dispatching Algorithm

The S-ED algorithm will be presented in this chapter referred to as SEconD. As mentioned in the previous chapter, SEconD incorporates wind speed as stochastic input and produce the resultant stochastic outputs.

2.1. Stochastic inputs and outputs

Hence in this dissertation, the wind speed for planning target time *t* will be denoted by $W(t)$ and its distribution will be denoted by $F_{W(t)}$; the dependence of the wind speed and its distribution on the target time *t* is explicitly emphasized. The wind power used in the ED models is a function of the wind speed, $P_{W(t)} = q\left(W(t)\right)$, thus the random fluctuation of the wind speed induces randomness into the wind power, thereby into the ED model [43] and [44]. The outputs of an ED model with *N* fossil fuel generators, include a vector of optimal powers $P_{ED} = \left(P_{1,ED}, \ldots, P_{N,ED}\right)$, the total operation cost *OC*, and the total transmission loss P_L. Due to the randomness of $P_{W(t)}$, all of the outputs of an ED model are subject to random fluctuations.

The following relationship between the wind power at the planning time, $P_{W(t)}$ and $W(t)$ is commonly used:

$$P_{W(t)} = q\left(W(t)\right) = \begin{cases} 0 & W(t) \le V_{CI} \text{ or } W(t) \ge V_{CO} \\ \dfrac{\left(W(t) - V_{CI}\right)W_r}{V_r - V_{CI}} & V_{CI} \le W(t) \le V_r \\ W_r & V_r \le W(t) \le V_{CO} \end{cases} \tag{2-1}$$

where V_{CI}, V_r and V_{CO} are cut-in, rated and cut-out wind speeds respectively, and W_r is the rated wind power.

The probability distribution of the wind power at the planning time, $P_{W(t)}$, consists of a continuous part over the interval (0, Wr) with the probability density function (PDF):

$$g\left(P_{W(t)}\right)=\frac{V_r-V_{CI}}{W_r} f_{W(t)}\left(P_{W(t)}\right), 0 \le P_{W(t)} \le w_r \tag{2-2}$$

where $f_{W(t)}$ is the PDF of $F_{W(t)}$, and two probability atoms at the end points of the interval given by

$$\Pr\left(P_{W(t)}=0\right)=1+F_{W(t)}\left(V_{CI}\right)-F_{W(t)}\left(V_{CO}\right), \tag{2-3}$$

$$\Pr\left(P_{W(t)}=w_r\right)=F_{W(t)}\left(V_{CO}\right)-F_{W(t)}\left(V_{CI}\right). \tag{2-4}$$

The time-variant random input, $P_{W(t)}$, of the ED model produces time-variant random fossil fuel power vector, $P_{ED}(t)=P_{ED}\left(W(t)\right)$.

The inputs of the proposed SEconD algorithm are Monte Carlo samples $W^1(t),\ldots,W^M(t)$ from a wind speed distribution. For each simulated wind speed, $W^h(t), h=1,\ldots,M$, an of outcome $P^h_{W(t)}=g\left(W^h(t)\right), h=1,...,M$ is computed using (2-1), which provide a random sample of size M from the probability distribution of the wind power (2-2), (2-3), and (2-4).

For each simulated input of the wind power, $P^h_{W(t)}=g\left(W^h(t)\right), h=1,...,M$ the ED model uses the power demand, fuel cost coefficients, transmission loss coefficients, and the system constraints as inputs. The ED model produces $P^h(t)=\left(P_1(t),\ldots,P_N(t)\right)^h$, $OC^h(t)$, and $P^h_L(t)$. Upon using the entire set of simulated samples of the wind power, $P^h_{W(t)}=g\left(W^h(t)\right), h=1,...,M$, the ED model provides Monte Carlo samples of size M from the probability distributions of the optimal random

vector $P_{ED}(t)$, the operation cost $OC_{ED}(t)$, and the transmission loss $P_{ED,L}(t)$. These samples are made available through simulations where mathematical functional relationships between the inputs and the outputs of an ED model are not available in closed-forms.

The Monte Carlo samples can be used to estimate distributional functions such as PDFs, cumulative probability distributions, reliability functions, and summary measures such the mean, standard deviation, median, percentiles, prediction intervals, correlation coefficients, and scatter plots. The algorithm is capable of generating a large number of samples, *M*, such that by the Law of Large Numbers, the distributions of $P_{ED}(t) = \left(P_{ED,1}(t), \ldots, P_{ED,N}(t)\right)$, $OC_{ED}(t)$, and $P_{ED,L}(t)$ can be reliably estimated.

The estimates of the entire distributions of the optimal outputs allow distributional comparison in their entirety such as stochastic ordering [45].

2.1.1. Stochastic Order

Various methods are available for comparing the distributions of two random variables [57]. One such a method used in this dissertation is the comparison of the reliability (survival) function,

$$R(x) = \Pr(X > x). \tag{2-5}$$

A random variable X_1 with reliability function R_1 is said to be stochastically smaller than or equal to random variable X_2 with reliability function R_2 if

$$R_1(x) \leq R_2(x), \text{ for every } x. \tag{2-6}$$

When this relationship holds, it is usually said that X_2 stochastically dominates X_1. For example, when X_i represents the optimal power generated by generator *i* and Eq. (2-6) holds, the probability that the optimal power generation by generator 2 exceeds any given value *x* MW is higher than the power generation by generator 1. In this case the stochastic dominance has a positive connotation.

2.2. Wind speed simulation model

For simulating wind speed data, we assume that a forecast, $W_f(t)$, for the target time t is available. (Developing the forecast is beyond the scope of this chapter and we refer the reader to [46]-[51] and [75]-[77]). We assume that $W_f(t)$ is a nonstochastic function of t, which provides the most likely wind speed for the target time. That is, $W_f(t)$ is the mode of the wind speed PDF, $f_W(t)$. This assumption corresponds to assuming that the actual wind speed of the target time will be $W_f(t)$ subject to a random error whose most likely value is zero. The following model combines the available forecast and the error:

$$W(t) = W_f(t) + m(t)\varepsilon . \tag{2-7}$$

where ε is a random variable with probability distribution F_ε and $m(t) > 0$ is a tuning parameter for determining the scale of the wind speed distribution $F_{W(t)}$ such that $W_f(t)$ is the mode of $f_W(t)$ and the mode of f_ε is at zero. Note that $m(t)$ is the ratio of the scale parameters of the wind speed and error distributions.

In forecasting literature it is common to assume that the error distribution F_ε is normal [52]. For such models, the mean and the median of the error distribution are also zero, the wind speed distribution $f_W(t)$ is normal with the mean and median equal to $W_f(t)$, and the standard deviation equal to the scale parameter, $\sigma_{W(t)} = m(t)\sigma_\varepsilon$. However, in the power literature, often it is assumed that the wind speed distribution is a Weibull, [44]-[45],[47]-[51], [53]-[55]. For the Weibull model with mode, $W_f(t)$, the mean, median, and the mode are different, and the corresponding error distribution is a three-parameter Weibull distribution with the following PDF:

$$f(\varepsilon)=\frac{k}{c}\left(\frac{\varepsilon-\tau}{c}\right)^{k-1} e^{-\left(\frac{\varepsilon-\tau}{c}\right)^{k}},\ \varepsilon\geq\tau, c>0, k>0, \tag{2-8}$$

where k is the shape parameter, c is the scale parameter, τ is referred to as the threshold parameter, and $\Pr(\varepsilon(t)\geq\tau)=1$ [56], [57]. Model (2-8) implies that $f_{W(t)}$ has the same shape parameter as f_{ε}, but its mode and scale are different from those for f_{ε}.

The two-parameter Weibull PDF for the wind speed is found by (2-7) and (2-8) with $\tau=0$, the shape parameter, k, and the scale parameter $c_{W(t)}=m(t)c$. Using the formula for the mode of the Weibull distribution relates the model parameters as follows:

$$W_f(t)=c_{W(t)}=m(t)c\left(\frac{k-1}{k}\right)^{\frac{1}{k}},\ k>1. \tag{2-9}$$

The forecast $W_f(t)$ induces the dynamic of time into mode and scale of the wind speed distribution, hence $F_{W(t)}$ is time-variant. The error distribution does not depend on the specific forecast for time t. Model (2-7) enables us to simulate the wind speed of the target planning time $W^h(t), h=1,...,M$ through the static distribution F_{ε} instead of the time-variant distribution $F_{w(t)}$.

We illustrate application of (2-7) using a given forecast $W_f(t)$ = 8 m/s for 9 a.m. for a wind farm in Bayfield, Wisconsin.

2.2.1. Illustration of wind speed data simulation

Although we assume that the distribution of the forecast error is available and forecasting is not the purpose of this chapter, we give an example to illustrate how data from the time-variant distribution of the wind speed can be simulated through a model such as (2-7).

The website www.windguru.cz/int/ provides data for three hour time sets for many locations in the North America. Model (2-7) illustrated using the wind speed profile of Bayfield-Wisconsin-USA for 9 a.m. January 2011-January 2013. Figure 2.1 shows the wind speed profile $W_f(t)$ which has non-random pattern including cyclic trend (estimates of $W_f(t)$) colored in red. Also Figure 2.2 depicts the autocorrelation function of the wind speed data indicates high autocorrelation between several lags.

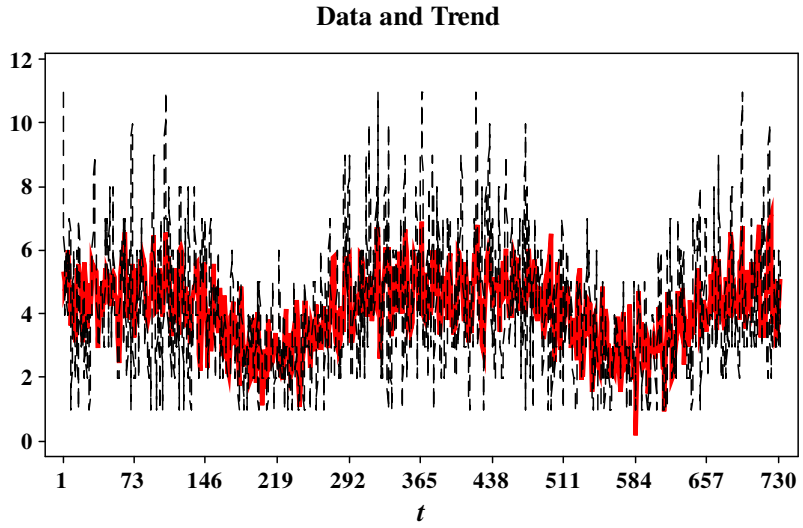

Figure 2-1 Time series plots of the wind speed and time trend.

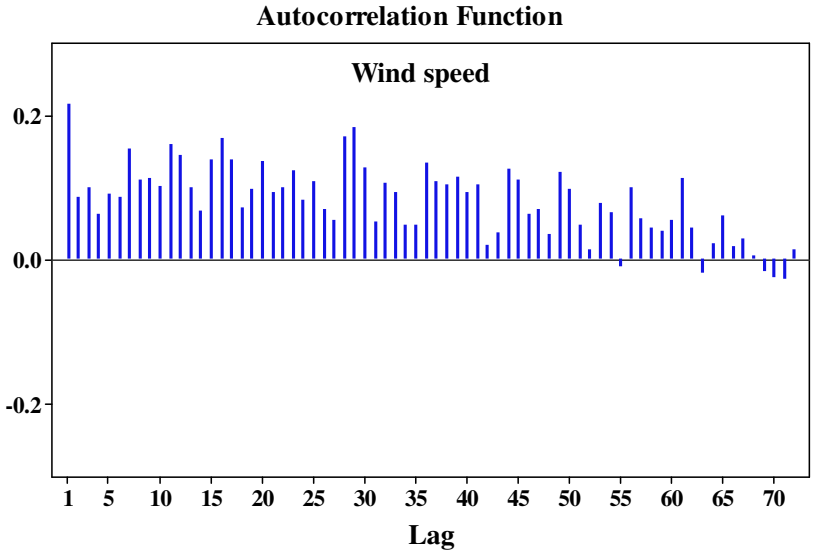

Figure 2-2 Autocorrelation function of the wind speed data

Next we develop an empirical version of (2-7) such that the residuals (estimates of $\varepsilon(t)$) pass an autocorrelation test [58]. A common method for analyzing a time series is based on the sine and cosine waves with different frequencies. The Fourier transform decomposition of the series W_t is:

$$W_t = \frac{a_0}{2} + \sum_{k=1}^{K}\left[a_k \cos(\omega_k t) + b_k \sin(\omega_k t)\right]. \quad (2\text{-}10)$$

where

t is the time subscript, $t = 1, 2, \ldots, n$

W_t are the data

n is the number of frequencies in the Fourier decomposition: $K = n/2$ if n is even; $K = (n-1)/2$ if n is odd

a_0 is the mean term: $a_0 = 2\overline{W}$

a_k are the cosine coefficients:

b_k are the sine coefficients

ω_k are the Fourier frequencies: $\omega_k = \frac{2\pi k}{n}$

The contribution of k^{th} term to series can be evaluated through periodogram. The amplitude of the periodogram, J_k, is defined as follows:

$$J_k = \frac{n}{2}\left(a_k^2 + b_k^2\right). \quad (2\text{-}11)$$

Therefore, in the example, we decomposed the wind speed data with a finite Fourier transform having eighteen pairs of sine and cosine which have the $J_k > 25$. We chose the lowest number of pair of Fourier terms with larger J_k while residuals passed the autocorrelation test. Also it should be noted that the model we considered with eighteen pairs of sine and cosine is not the only model. While a model passes the autocorrelation test can be used to decompose wind speed data. Figure

2.3 shows the periodogram of the detrended wind speed data. This figure confirms that the total number Fourier terms having $J_k > 25$ are eighteen.

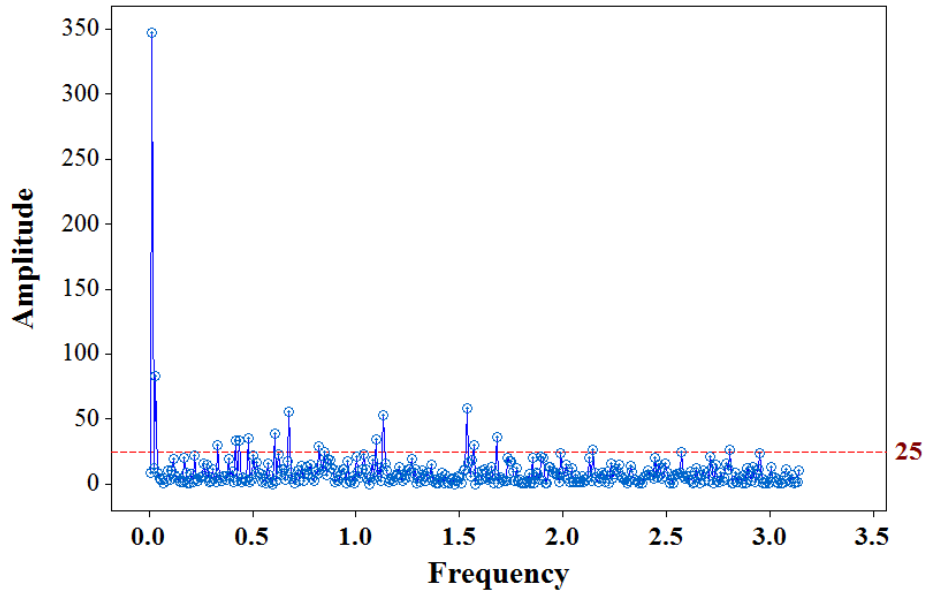

Figure 2-3 Periodogram of the detrended wind speed data with the amplitude reference line

Figure 2.4 shows the time series plot of the residuals which passes the stationary test. The Autocorrelation function of the residuals with 5% significant band depicted in Figure 2.5.

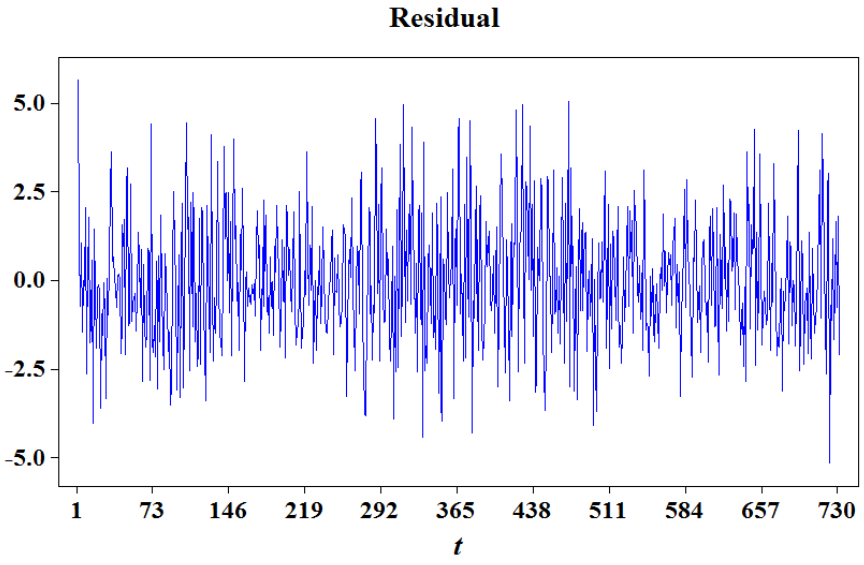

Figure 2-4 Time series plots of the residuals.

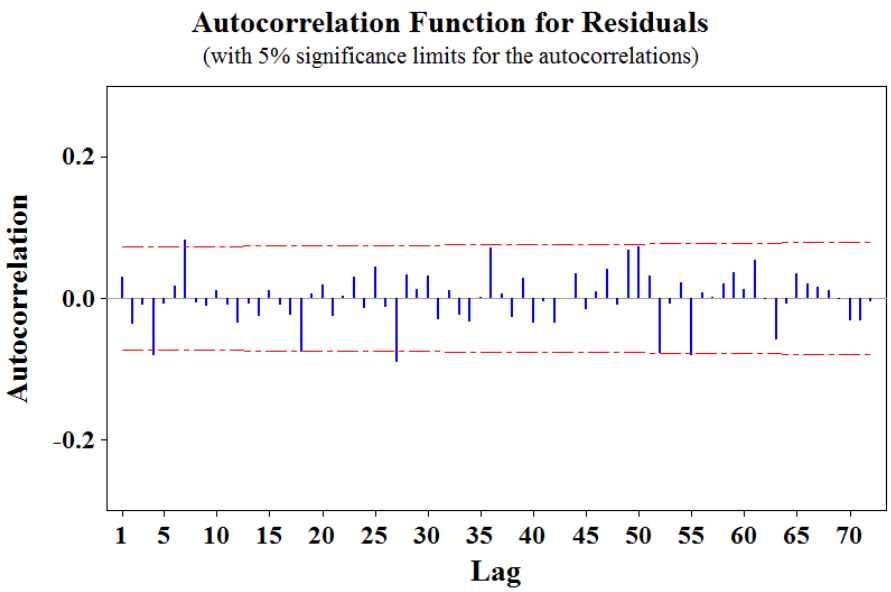

Figure 2-5 Autocorrelation function of the residuals

Figure 2.6 shows the histogram of the residuals superimposed by the PDF of the three-parameter Weibull distribution. The parameters of the Weibull model are shown on the graph. It should be noted that the formula of the significant band for autocorrelation function is based on the assumption that the distribution of the epsilons is normal. For the distribution of the residuals, the Anderson-Darling test for normal model gives AD=1.108 (P-value=0.011) and for a three-parameter Weibull gives AD=0.725 (P-value = 0.040). In fact the normal and Weibull distributions that fit the residuals are very similar. Since the AD test does not reject the normal assumption at 1% level we can use the band and infer that the epsilons are independent. For the simulations we can use either of these two models, however, according to the AD, the fit of the Weibull is better than the normal, so we use Weibull. Moreover, the Weibull model for the error produces positive wind speed data with probability one, however, this not the case for the normal model [59].

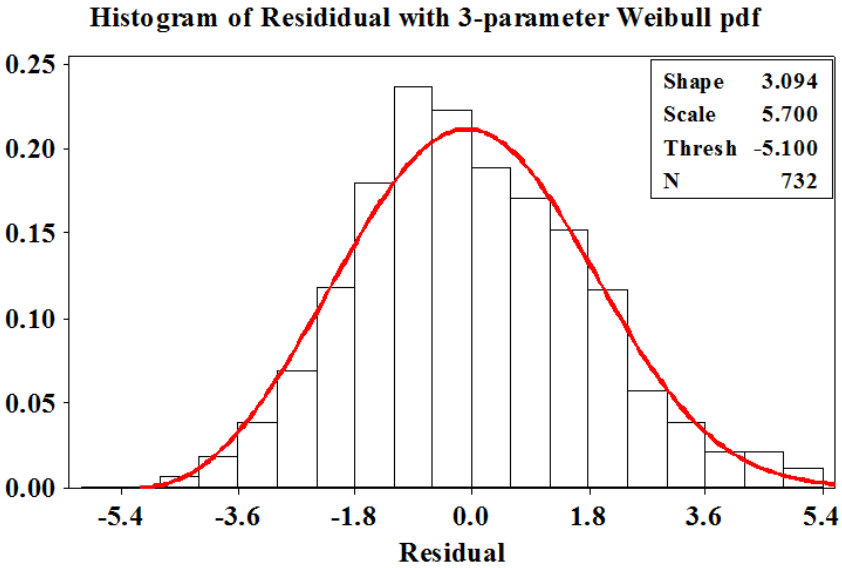

Figure 2-6 Histogram of the residuals and three-parameter Weibull PDF.

The error distribution is found through a model selection exercise described above. The forecast value combined with the error distribution via (2-8) gives the Weibull distribution shown in Figure 2.7 for simulating the wind speed data.

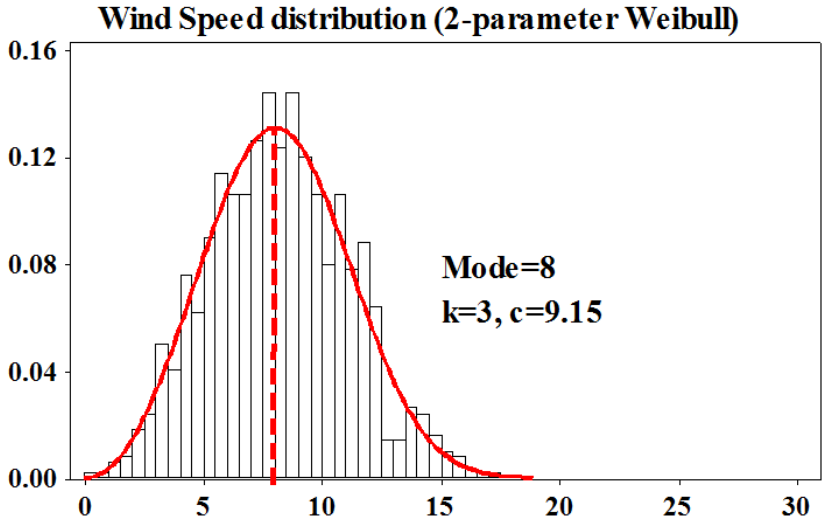

Figure 2-7 Histogram of the simulated wind speed data superimposed by the Weibull PDF.

The Weibull PDF shown in Figure 2.7 is obtained as follows. Using $W_f(t) = 8$ [m/s] and the shape parameter of the error distribution $k = 3$ in (2-9) gives the scale parameter of the Weibull distribution for the wind speed as $c = 9.15$. In this example, the scale tuning parameter is

$m(t)=1.605$. Figure 2.7 also shows the histogram of a sample M = 1,000 simulated wind speed data the Weibull PDF.

2.3. ED model

In this example we use the simulated wind power samples $P_W(t)^h, h=1,...,M$ as inputs into the conventional ED model formulated as follows:

$$\min_P OC=\sum_{i=1}^{N} a_i P_i^2+b_i P_i+c_i \tag{2-12}$$

where $P=(P_1,\ldots,P_N)$ is subject to operation ranges

$$P_i^{\min} \le P_i \le P_i^{\max} \tag{2-13}$$

and the power supply-demand balance equation is:

$$P_D+P_L-\sum_{i=1}^{N} P_i=0 \tag{2-14}$$

where P_{N+1} represents the simulated wind power samples, $P_{N+1}\equiv P_{W(t)}^h, h=1,...,M$, a_i, b_i and c_i are power cost coefficients and P_L is the transmission loss given by *Kron's* formula:

$$P_L=\sum_{i=1}^{N+1}\sum_{j=1}^{N+1} P_i B_{ij} P_j+\sum_{i=1}^{N+1} B_{i0} P_i+B_0 \tag{2-15}$$

For each simulated sample point $W^h(t)$, the software computes the wind power, $P_{W(t)}^h$. In this illustration, the inequality bounds in (2-1) are set as V_{CI}= 5 m/s, V_r = 14 m/s, V_{CO} = 17 m/s, and W_r= 30 MW. The wind power samples $P_{W(t)}^h, h=1,...,M$ in (2-3) and (2-4) carry the time-variant randomness of the wind speed into the conventional ED model. This section illustrates the

implementation and results of the SEconD algorithm for a wind-penetrated system consisting of one wind farm, and $N = 3$ fossil fuel power generators for $P_D = 210$ MW power demand. Table 2.1 shows the operation ranges, cost coefficients, and the transmission loss matrices used in this example.

Table 2-1 Operation ranges and power cost coefficients.

Generator	Range		Power Cost Coefficients		
	P_{min}	P_{max}	a_i	b_i	c_i
1	54	110	0.00633	11.669	213.1
2	50	120	0.00889	10.333	200.0
3	47	100	0.00741	10.833	240.0

$$\left[B_{ij}\right] = \begin{bmatrix} 0.05370 & 0.00453 & 0.00207 & 0.04507 \\ 0.00453 & 0.06760 & 0.00953 & 0.00507 \\ 0.00207 & 0.00953 & 0.05210 & 0.00901 \\ 0.04507 & 0.00507 & 0.00901 & 0.2940 \end{bmatrix}$$

$$\left[B'_{oj}\right] = \begin{bmatrix} 0.00330 & 0.07660 & 0.00342 & 0.01890 \end{bmatrix}$$

$$\left[B_0\right] = \begin{bmatrix} 0.040357 \end{bmatrix}$$

The unit of P_i is MW, and the units of a_i, b_i and c_i are respectively \$/MW^2 h, \$/MWh, and \$/h. Thus, the unit of OC is \$/h. This section illustrates the implementation and results of the SEconD algorithm for a wind-penetrated system consisting of one wind farm, and $N = 3$ fossil fuel power generators for $P_D = 210$ MW power demand.

2.3.1. Probability distribution of optimal outputs

Table 2.2 gives the summary statistics for the inputs (simulated wind speed data and wind power generation) and the optimal outputs of the three generators, the operation cost, and the transmission loss. Note that the ranges of the optimal outputs of the generators (between the minimum and maximum) are subintervals of their operation ranges. Comparing the means or the medians provides the following information: on average, the optimal output of generator 3 is higher than the optimal output generator 2, which in turn is higher than the optimal output generator 1.

Table 2-2 Summary statistics.

	Min	Max	Median	Mean	Std.
WS	0.00	18.00	8.18	8.26	2.84
P_w	0.00	30.00	10.59	11.31	8.42
P_1	56.39	59.51	57.43	57.50	1.15
P_2	71.06	76.22	75.18	74.78	1.31
P_3	60.61	83.53	76.19	75.21	6.47
OC	2744.00	3112.70	2976.40	2972.90	103.80
P_L	9.25	9.48	9.31	9.33	0.06

Figure 2.8 shows the PDFs of the optimal output powers of the three fossil fuel generators along with the segments (2-2), (2-3), and (2-4) of the wind power distribution, all computed from the Monte Carlo samples. The two probability atoms shown in the PDF of P_W represent $\Pr(P_W = 0)$ and $\Pr(P_W = W_r)$, where $W_r = 30$ MW. As seen in this figure, the distribution of P_1 is concentrated near the lower bound of operation specified in Table 2.1. The distribution of P_2 is concentrated in the middle of its bounds, and the distribution of P_3 is less concentrated than the other two. The PDF plots indicate higher uncertainty about the power generation for the third generator. Consequently, the system operator may consider setting aside more reserve for this generator. The most probable values of power outputs P_1, P_2, and P_3 are around the modes of the distribution 59.51, 76.22, and 83.53 MW, respectively. However, due to the lack of concentration, the probability of an outcome of P_3 to be near the mode of the distribution is substantially less than that for P_1 and P_2.

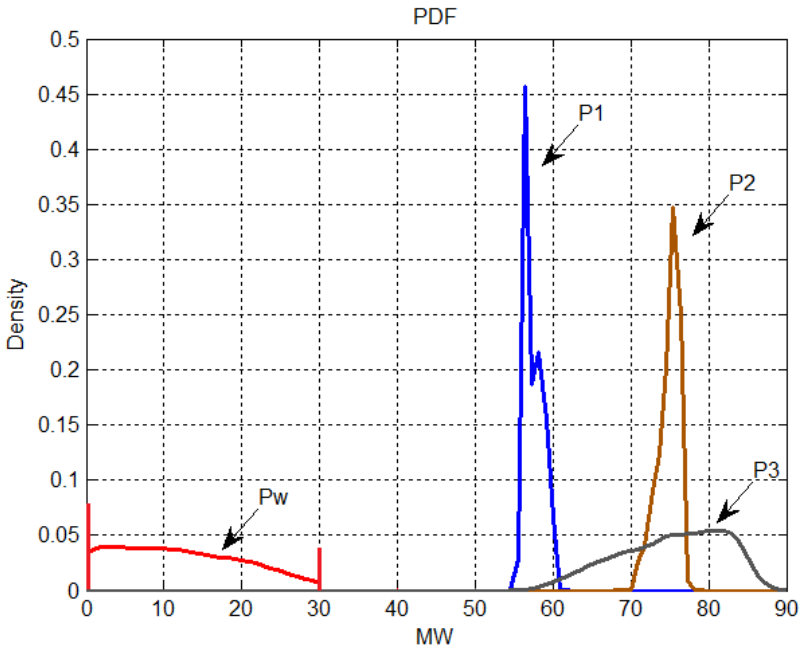

Figure 2-8 PDFs of the wind power and optimal output power generation.

More complete comparisons of the optimal power outputs are provided by their reliability functions shown in Figure 2.9. These plots provide descriptive measures such as percentiles. In this figure for any probability shown on the vertical axis, the corresponding upper percentile of the optimal output of a generator can be easily found on the horizontal axis. For example, the medians are the 50^{th} percentiles given by the points corresponding to 0.5 on the vertical axis of the reliability plots. As another example, Figure 2.9 shows that the upper 10^{th} percentile of P_3 is 83 MW, which is higher than the upper 10^{th} percentiles of P_1 and P_2. In fact, any upper percentile of P_3 is higher than the corresponding upper percentile of P_1. Accordingly, P_3 is said to be stochastically larger than P_1. Similarly, P_2 stochastically dominates P_1. These results imply that the probability that the optimal outputs of P_2 and P_3 exceed any given value, x MW on the horizontal axis, is higher for generators 2 and 3 than for generator 1. Alternatively, it can be seen that for a given probability, outputs of P_2 and P_3 are higher than the output of P_1. However, there is no stochastic dominance

between P_2 and P_3. For outputs less than 74 MW, P_2 dominates P_3 and vice versa for outputs more than 74 MW.

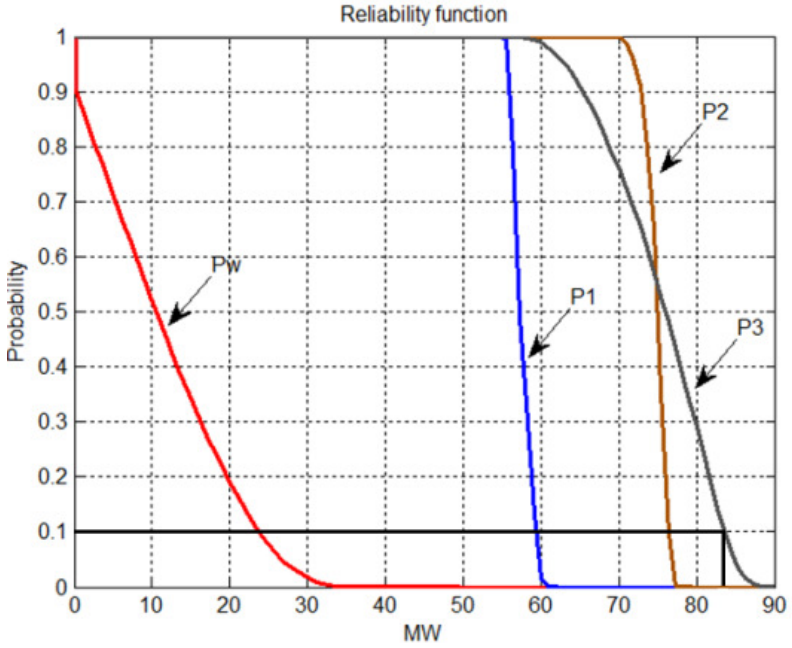

Figure 2-9 Reliability functions of the wind power and optimal output power generation.

2.3.2. Relationship between outputs and wind speed

The SEconD algorithm provides information for examining relationships between the wind speed and the optimal outputs of the ED model. Figure 2.10 shows the scatter plots of the optimal outputs of three fossil fuel generators as well as the wind power versus wind speed. These plots show that the required outputs of the fossil fuel generators decrease as the wind speed increases within the V_{CI} and V_r. Because of the cost parameters and transmission loss coefficients, P_3 decreases relatively faster than the other two outputs. The plots also show the maximum of the three fossil fuel generators when the wind speed is below V_{CI} or above V_{CO}. For any wind speed, the outputs of generators 2 and 3 are higher than the output of generator 1. This relationship is due to the

stochastic dominance noted in Figure 2.9. However, in Figure 2.10, the outputs of P_2 and P_3 cross at a point near 74 MW, the same point where their reliability functions crossed in Figure 2.9.

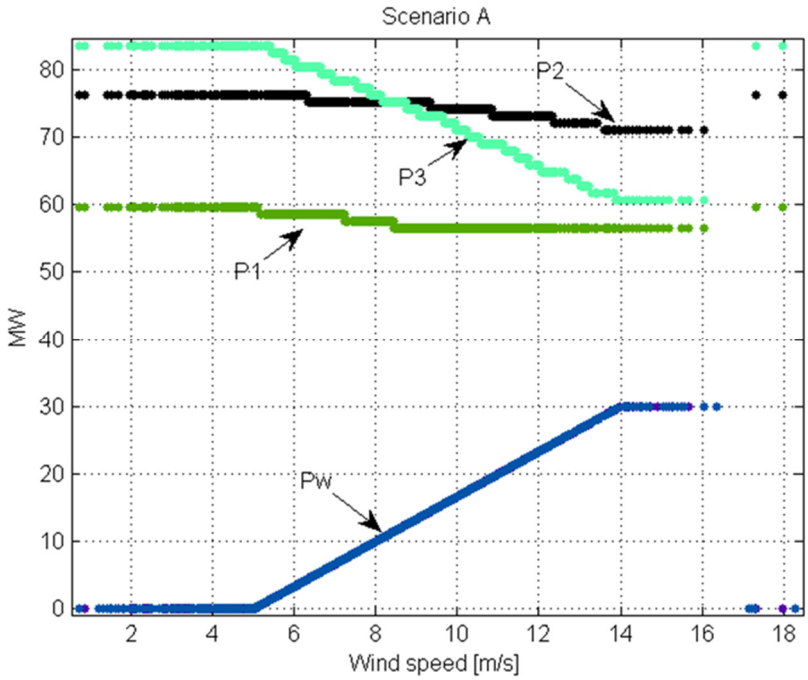

Figure 2-10 Scatter plots of the optimal outputs of three fossil fuel generators and wind power versus wind speed.

Table 2.3 gives the correlation coefficients between the optimal outputs of the three generators along with the correlation with the wind power generation and the wind speed. The correlations among the variables are high. The correlations between wind power and the optimal fossil fuel generators' outputs are negative, most strongly with P_3. It should be noted that the correlation between wind speed and wind power is strong, but not perfect due to the cut-in and cut-out wind speeds.

Table 2-3 Correlation coefficients.

	WS	Pw	P_1	P_2	P_3
WS	1				
Pw	0.96	1			
P_1	-0.81	-0.86	1		
P_2	-0.91	-0.95	0.71	1	
P_3	-0.95	-0.99	0.84	0.95	1

2.3.3. Planning applications

2.3.3.1. Prediction intervals for optimal outputs

The outputs of the SEconD algorithm presented in the preceding section enables a system operator to perform various analyses.

For predictive purposes, ranges corresponding to a specified probability (confidence level) can be easily found from the reliability plots in Figure 2.9. There are an infinite number of ranges that will produce the same probability. In the SEconD algorithm, the uncertain parameter is wind speed. In order to find a more accurate solution with the least margin of error for a given probability, the shortest range of a wind speed at a given probability must be found. Statistical algorithms give the intervals with the shortest margin of errors called *highest probability density* (HPD) intervals [60], [61]. These intervals are shown in the second column of Table 2.4. The two sets of intervals in Table 2.4 provide choices for the system operator to predict the output of each generator with the confidence levels shown in the table. It should be emphasized that these intervals are for each quantity individually.

Table 2-4 Prediction intervals.

	80% Range for each variable	Intervals corresponding to 80% range for WS	Intervals corresponding to 65-75MW range of P_3
WS	(3.7, 12.3), 80%	(3.7, 12.3), 80%	(8.0, 12.0), 15%
P_w	(0.0, 26.4), 80%	(0.0, 26.4), 80%	(10.0, 25.2), 35%
P_1	(56.4, 58.5), 80%	(56.4, 57.5), 63%	(56.4, 57.8), 78%
P_2	(73.1, 76.2), 80%	(72.1, 75.9), 73%	(72.1, 75.0), 60%
P_3	(65.8, 83.5), 80%	(65.8, 82.5), 76%	(65.0, 75.0), 35%

Once the distributions of the optimal results of the economic dispatch are obtained, the system operator can perform many useful analyses without further runs of the SEconD algorithm. Two examples using the combination of Figure 2.9 and Figure 2.10 are as follows.

1. Suppose that the system operator wishes to obtain the economic dispatching on the fossil fuel power plants based on a range of the wind speed with a given probability, 80% for instance. The ranges for the optimal power outputs corresponding to the shortest 80% range for the wind speed (shown in Table 2.4) can be easily found in Figure 2.10. These ranges are shown in the third column of Table 2.4 along with their probabilities found using the plots of the reliability functions in Figure 2.9.
2. Suppose that the system operator wishes to regulate the output of one of the fossil fuel power plants. For example, consider the 65-75 MW subset of the optimal range of operation 60.61-83.53 MW of P_3 given in Table 2.2. This control can be done with only 15% probability given by the reliability function of P_3 shown in Figure 2.9. Using Figure 2.10, the system operator finds the optimal operation ranges of other energy sources corresponding to 65-75 MW range of P_3. These ranges are shown in the last column of Table 2.4 with their probabilities found using the plots of the reliability functions in Figure 2.9.

2.3.3.2. Sensitivity analysis

The system operator can also perform "what if?" analyses of changes in the input parameters such as the costs of fuels (power cost coefficients), and the capacity limits of the generators (generators' operation ranges). For example, consider a scenario where the coefficients for the first generator in Table 2.1 are decreased by 5%, the second unit remains unchanged, and that of the third unit are increased by 3%. What are the effects of these changes on results reported using the parameters in Table 2.1. We refer to the original scenario as "A" and the altered scenario as "B". The fossil fuel generators are complicated functions of the cost coefficients. Thus, the implied changes in the distributions of the outputs must be determined through simulations using the new parameters. The

upper panel of Figure 2.11 shows the PDF plots of the distributions of the optimal output powers of the three fossil fuel generators, under scenario B. As seen in this figure, the effects of the perturbation are the shift of the distribution of P_3 closer to its lower bound and the shift of the distribution of P_1 closer to its upper bound. The PDF plots indicate higher uncertainty about the power generation for the third generator again. Consequently, the system operator may consider setting aside more reserve for this generator. The most probable values of power outputs P_1, P_2, and P_3 are around the modes of the distribution 81.8, 74.5, and 51 MW, respectively. However, due to the lack of concentration, the probability of an outcome of P_3 to be near the mode of the distribution is substantially less than that for P_1 and P_2. The reliability plots for scenario B are shown in the lower panel of Figure 2.11. These reliability plots show the stochastic dominance of P_1 over P_2 and P_3 as well as P_2 and P_3. Figure 2.12 shows the PDFs and reliability functions of the distributions of the minimal cost of power production for both scenarios. The PDFs for the two scenarios cross a few times, and thus do not provide a definite comparison for the two operation costs. However, the plots of the reliability functions provide a definite comparison: The optimal operation cost of scenario A is stochastically larger than that of scenario B. That is, for any given value, x \$/h, the probability that the optimal operation cost exceeds x is higher for scenario A than for scenario B. For instance, the probability that the optimal operation cost exceeds 3050 \$/h is about 30% and 27% for scenarios A and B, respectively. Consequently, scenario B is more cost effective than A. As a further example, suppose that the system operator wishes to compare the operation costs based on a range of the wind speed with a given probability. Also the 80% range of the minimum operation costs for scenarios A and B are (2804, 3112) and (2795, 3095) \$/h, respectively.

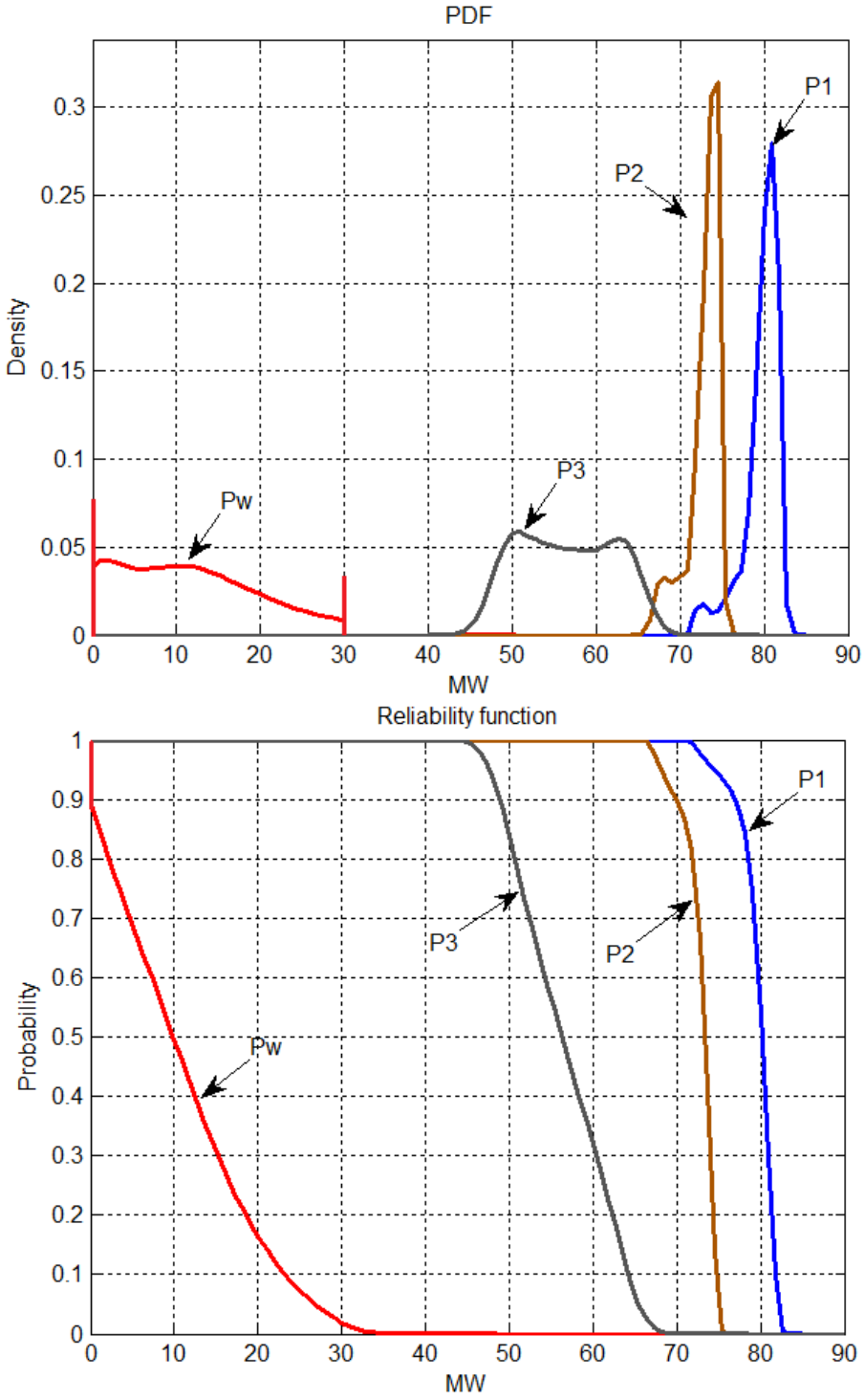

Figure 2-11 Distributions of the wind power and optimal output power generation under scenario B.

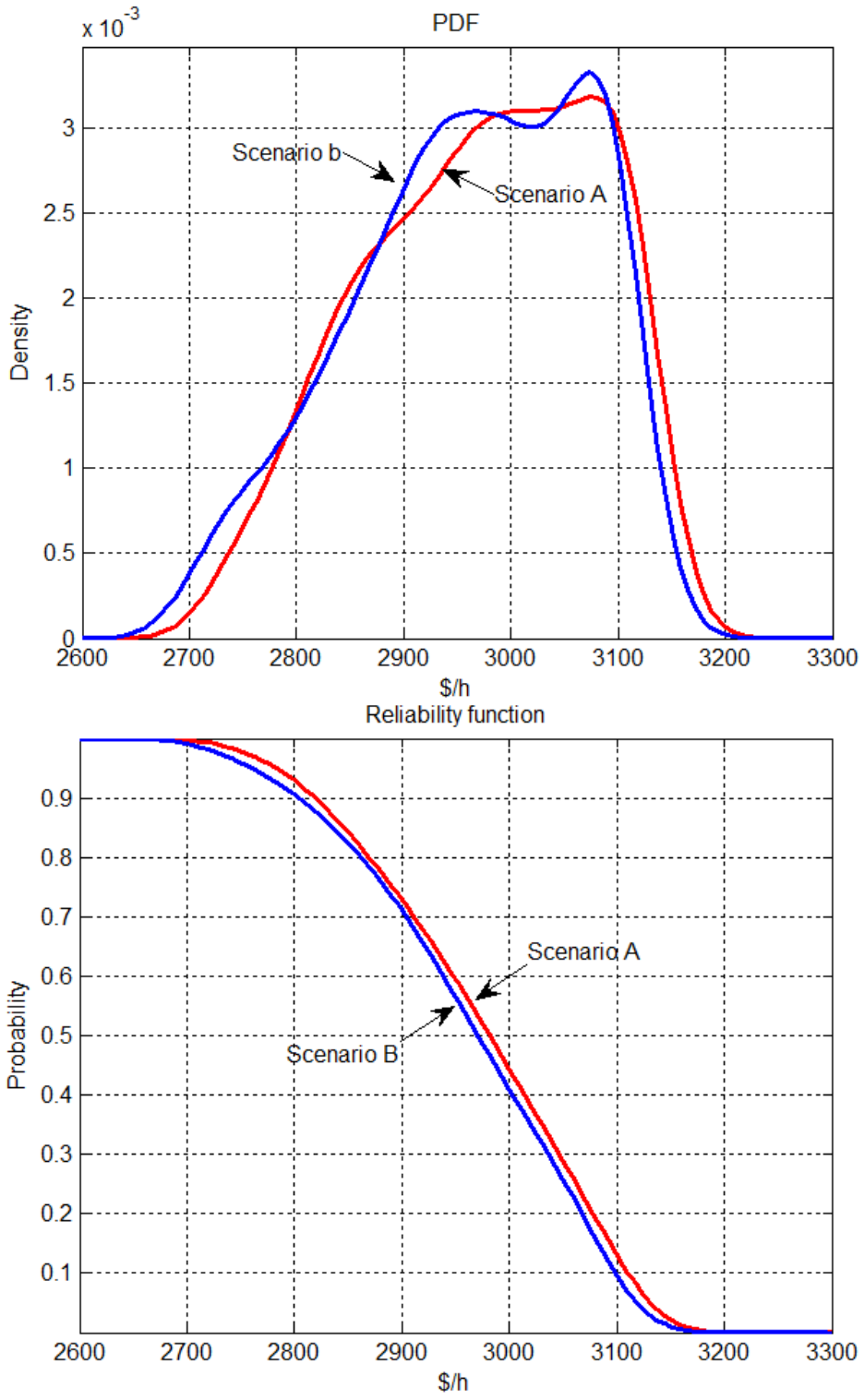

Figure 2-12 Distributions of the total costs for scenarios A and B.

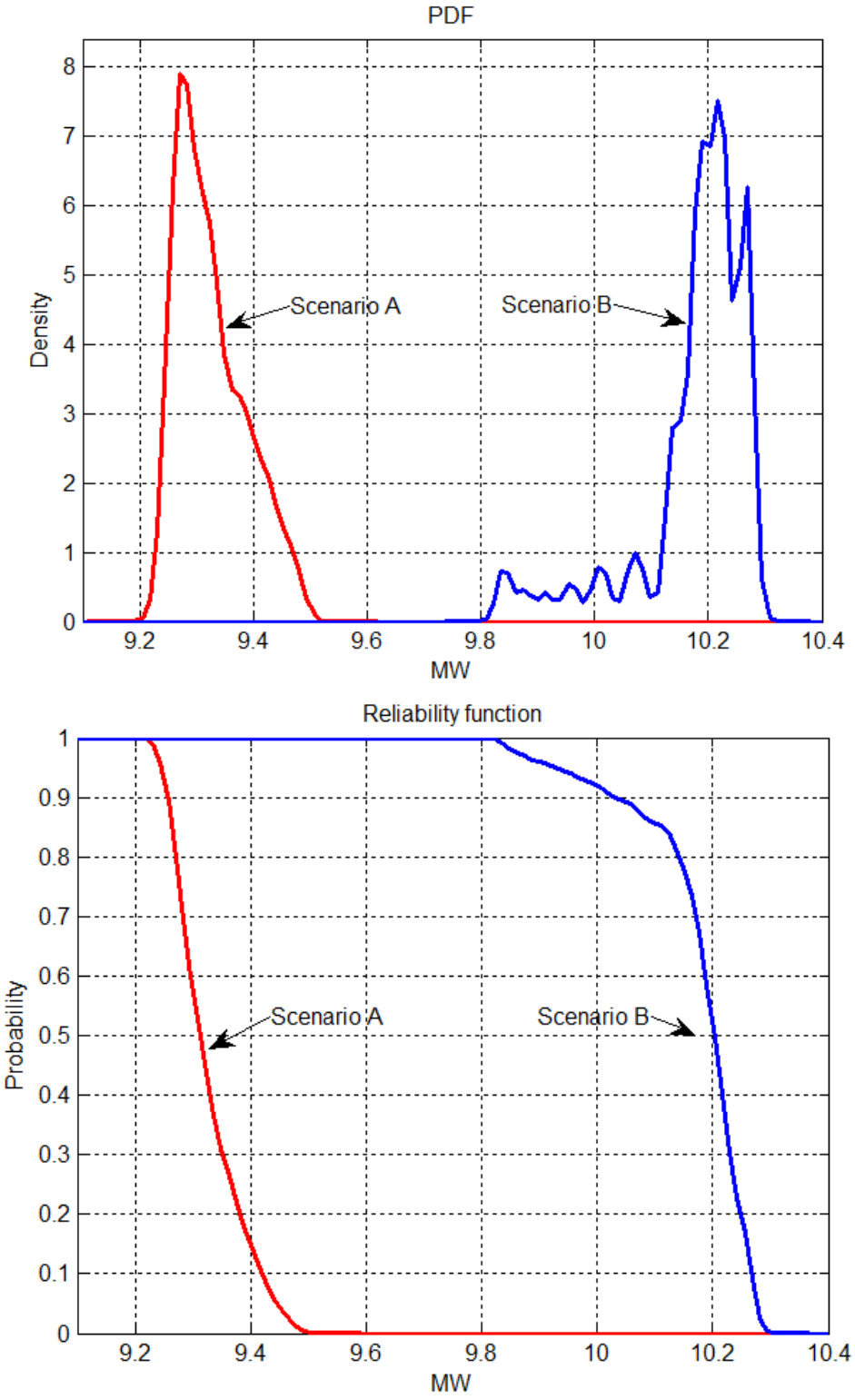

Figure 2-13 Distributions of the transmission losses for scenarios A and B.

Figure 2.13 shows the PDFs and reliability functions of the distributions of the system transmission loss for both scenarios. The plots of PDFs indicate that the most likely MW values of the

transmission losses are almost 9.27 and 10.21 for both scenarios A and B, respectively. However, the plots of the reliability functions provide a more complete comparison. These plots show the stochastic dominance of the transmission loss of scenario B over A.

3. Chapter 3

Bayesian Models for Wind-Penetrated Power Systems

In statistics, and many related fields dealing with data analysis, inclusion of uncertainty about the parameters of a probability distribution is very common. This approach, known as Bayesian, considers the distributional parameters as unknown or partially known quantities varying according to some probability distributions. In the Bayesian approach, sample data provide information about the distributions of the model parameters and prediction of future outcomes [62]. The Bayesian methods have astonishingly grown during the last two decades due to availability of Markov Chain Monte Carlo (MCMC) algorithms and software packages which facilitate estimation of the posterior distributions of the parameters and predictive distributions [63] and [64]. However, thus far, the uncertainty about the probability distribution of the wind speed has not been included in the context of wind-penetrated power systems models. Attention to this issue is crucial in order to avoid assuming perfect knowledge of the distributional parameters which, on average, would lead to the use of unduly optimistic and insufficiently disperse distributions for the wind speed for the planning time.

3.1. Proposed method

This chapter will develop Bayesian stochastic models which incorporate the uncertainty about the wind speed distributional parameters in order to use a Bayesian predictive distribution of the wind speed in the context of wind-penetrated power systems. More specifically, instead of using a wind speed distribution $f(w|\theta)$, where θ is a vector of parameters whose values are known or estimated, we will view θ as a random vector whose variation is according to a prior probability distribution

$p(\theta)$ which can be updated in light of data **D** into a posterior distribution $p(\theta|\mathbf{D})$ via the Bayes' rule as follows:

$$p(\theta|\mathbf{D}) \propto p(\mathbf{D}|\theta)p(\theta), \text{ "}\propto\text{" stands for proportional"} \quad (3\text{-}1)$$

where $p(\mathbf{D}|\theta)$ is the likelihood function obtained from the probability model of the data.

The Bayesian approach is prudent because of the inclusion of the uncertainty about the parameters in the model, which induces uncertainty about the wind speed distribution itself. The Bayesian prior predictive distribution of the wind speed is given by

$$f(w) = \int f(w|\theta)\, p(\theta)\, d\theta, \quad (3\text{-}2)$$

The posterior predictive distribution uses the same formula where $p(\theta)$ is replaced with $p(\theta|\mathbf{D})$.

$$f(w_{n+1})|\boldsymbol{D}) = \int f(w_{n+1}|\theta)\, p(\theta|\boldsymbol{D})\, d\theta, \quad (3\text{-}3)$$

where $f(w_{n+1}|\theta)$ is the conditional density function of w_{n+1}.

Bayesian methods are available for situations where there is a complete absence of knowledge about the parameters as well as when some partial information about the distribution of the parameters. In the case of the complete absence of knowledge about the parameter, several methods for developing non-informative priors have been suggested in the statistics literature [65] and for the case of partial information developing prior distribution based on the maximum entropy approach are used [65] and [66].

Two important consequences of a Bayesian stochastic model are as follows.

a) Inclusion of uncertainty about the parameters of the wind speed distribution results in using a more prudent predictive distribution for the wind speed. That means, on average, the predictive distribution is more disperse than the probability distributions

when the uncertainty about the parameters is ignored. Consequently, for example, for a range of the wind speed with a given probability, the range under the Bayesian predictive distribution is wider than that of ignoring the parameter uncertainty. Conversely, for a given probability, the range of the wind speed under the Bayesian predictive distribution is narrower than that of ignoring the parameter uncertainty.

b) The probability distributions of the parameters can be viewed in terms of the heterogeneity of the distributions of the wind speed over a wind farm. The wind speed distributions for various turbines in a farm may belong to the same family of models, such as the Weibull, and the model parameters of each turbine may vary randomly according to some probability distributions. The Bayesian predictive distribution aggregates the non-homogeneous distributions into a single distribution that captures the variation among the probability distributions of the wind speeds at the turbines' locations in a wind farm.

3.2. Example: Rayleigh model for data

The Rayleigh distribution is the simplest distribution commonly used to describe wind speeds [22], [67] and [68] because it only has a single model parameter *c*. Assume that the wind speed distribution is Rayleigh with the PDF,

$$f(w|c) = \left(\frac{2}{c}\right)\left(\frac{w}{c}\right)e^{-\left(\frac{w}{c}\right)^2}. \qquad (3\text{-}4)$$

Let $D=(\mathrm{w}_1,...,\mathrm{w}_n)$ denote the wind speed profile. Model (3-4) provides the following likelihood function under the independency assumption:

$$\begin{aligned} p(\mathbf{D}\mid c) &= f(w_1,\ldots,w_N \mid c) \\ &= f(w_1 \mid c) f(w_2 \mid c)\ldots f(w_N \mid c) \\ &= \left(\frac{2}{c}\right)^n e^{-\frac{\sum w_i^2}{c^2}} \prod \frac{w_i}{c}. \end{aligned} \tag{3-5}$$

For simplifying the illustration we re-parameterize $\theta = \frac{1}{c^2}$ and let $\sum w_i^2 = T_n$.

Because w_i are not a function of θ then,

$$L(\theta) \propto \theta^n e^{-\theta T_n} \tag{3-6}$$

3.2.1. Prior information

Consider the general problem of inferring a distribution for a parameter θ given some datum or data w_i. From Bayes' theorem, the posterior distribution is proportional to the product of the likelihood function $\theta \rightarrow p(w_i \mid \theta)$ and prior $p(\theta)$, $p(\theta \mid w) \propto L(\theta)\, p(\theta)$.

In the Bayesian analysis, if the posterior distributions $p(\theta|w_i)$ are in the same family as the prior distribution $p(\theta)$, the prior and posterior are then called conjugate distributions, and the prior is called a conjugate prior for the likelihood function. A conjugate prior is an algebraic convenience, giving a closed-form expression for the posterior; otherwise a difficult numerical integration may be necessary. Further, conjugate priors may give intuition by more transparently showing how a likelihood function updates a prior distribution. All members of the exponential family have conjugate priors. The family of conjugate priors for (3-6) is gamma prior $G(\alpha,\beta)$. The Gamma PDF, expected value and variance are as follows:

$$G(\alpha,\beta) = \frac{\beta^\alpha}{\Gamma(\alpha)} \theta^{\alpha-1} e^{-\beta\theta}, \tag{3-7}$$

$$E(\theta)=\frac{\alpha}{\beta}, \tag{3-8}$$

$$Var(\theta)=\frac{\alpha}{\beta^2}. \tag{3-9}$$

From the Bayes' rule (1) gives the posterior:

$$\begin{aligned} p(\theta|T_n) &\propto \theta^{\alpha-1}e^{-\beta\theta}\theta^n e^{-\theta T_n} \\ &\propto \theta^{n+\alpha-1}e^{-(\beta+T_n)\theta} \end{aligned}. \tag{3-10}$$

That means the posterior distribution is also $G(n+\alpha,\beta+T_n)$ and under quadratic loss the Bayes' estimate is:

$$\tilde{\theta}=E(\theta|\boldsymbol{D})=\frac{n+\alpha}{\beta+T_n}, \tag{3-11}$$

and the posterior variance is:

$$Var(\theta|\boldsymbol{D})=\frac{n+\alpha}{(\beta+T_n)^2}, \tag{3-12}$$

The Bayesian predictive distribution of wind speed is:

$$\begin{aligned} f(w|D) &\propto \int \theta w e^{-\theta w^2}.\theta^{n+\alpha-1}e^{-(\beta+T_n)\theta}d\theta \\ &\propto w\int \theta^{n+\alpha+1-1}.e^{-(w^2+\beta+T_n)\theta}d\theta \\ &\propto w(w^2+\beta+T_n)^{-(\alpha+n)+1} \\ &\propto w\left(1+\frac{w^2}{\beta+T_n}\right)^{-(\alpha+n)+1} \end{aligned}. \tag{3-13}$$

Expression (3-14) is the kernel of the Pareto Type IV distribution, *P*(IV), with scale parameter (T_n + β)$^{1/2}$, shape parameter 2, and the tail index $\alpha+n$. The complete PDF of predictive wind speed is:

$$f\left(w|D\right)=2w\left(\alpha+n\right)\left\{\left(T_n+\beta\right)\left(1+\frac{w^2}{T_n+\beta}\right)\right\}^{-(\alpha+n+1)},w\geq 0. \tag{3-14}$$

The survival function, $\overline{F}\left(w|D\right)$, derived from $f(w|\boldsymbol{D})$ is:

$$\overline{F}\left(w|D\right)=\left(1+\frac{w^2}{\beta+T_n}\right)^{-(\alpha+n)}, \tag{3-15}$$

The mean of (3-14) is:

$$\begin{aligned} E\left(W\right) &= \frac{\left(T_n+\beta\right)^{1/2}.\Gamma\left(\alpha+n-2\right).\Gamma\left(3\right)}{\Gamma\left(\alpha+n\right)} \\ &= \frac{2\left(T_n+\beta\right)^{1/2}}{\left(\alpha+n-1\right)\left(\alpha+n-2\right)} \end{aligned}. \tag{3-16}$$

The 2-moment of (3-14) is:

$$\begin{aligned} E\left(W^2\right) &= \frac{\left(T_n+\beta\right)^{2}.\Gamma\left(\alpha+n-4\right).\Gamma\left(5\right)}{\Gamma\left(\alpha+n\right)} \\ &= \frac{24\left(T_n+\beta\right)^{2}}{\left(\alpha+n-1\right)\left(\alpha+n-2\right)\left(\alpha+n-3\right)\left(\alpha+n-4\right)} \end{aligned}. \tag{3-17}$$

The variance of (3-14) is:

$$\begin{aligned} Var\left(W\right) &= E\left(W^2\right)-E\left(W\right)^2 \\ &= \frac{24\left(T_n+\beta\right)^{2}}{\left(\alpha+n-1\right)\left(\alpha+n-2\right)\left(\alpha+n-3\right)\left(\alpha+n-4\right)}-\frac{4\left(T_n+\beta\right)}{\left(\alpha+n-1\right)^2\left(\alpha+n-2\right)^2} \end{aligned}. \tag{3-18}$$

The mean and variance of the distribution (3-14) exist because $2 < \alpha + n$.

Understanding how changes in the model inputs influence the outputs is a concern. In this chapter, we do the sensitivity analyses via changing the number of wind samples, parameters of prior distributions and class of prior distributions.

3.2.2. Simulations and results

In this section we simulate small wind speed samples, say n = 3, 5 and 20, from a known Rayleigh model without uncertainty with c = 9.24. Then we do posterior and predictive inferences under two informative priors on the parameter $\theta = 1/c^2$. We begin with presenting priors for θ as shown in Figure 3.1.

Prior 1, Gamma with $\alpha = 1$ and $\beta = 1$, $G(1,1)$, which represents an exponential distribution

Prior 2, Gamma with $\alpha = 10$ and $\beta = 10$, $G(10,10)$,

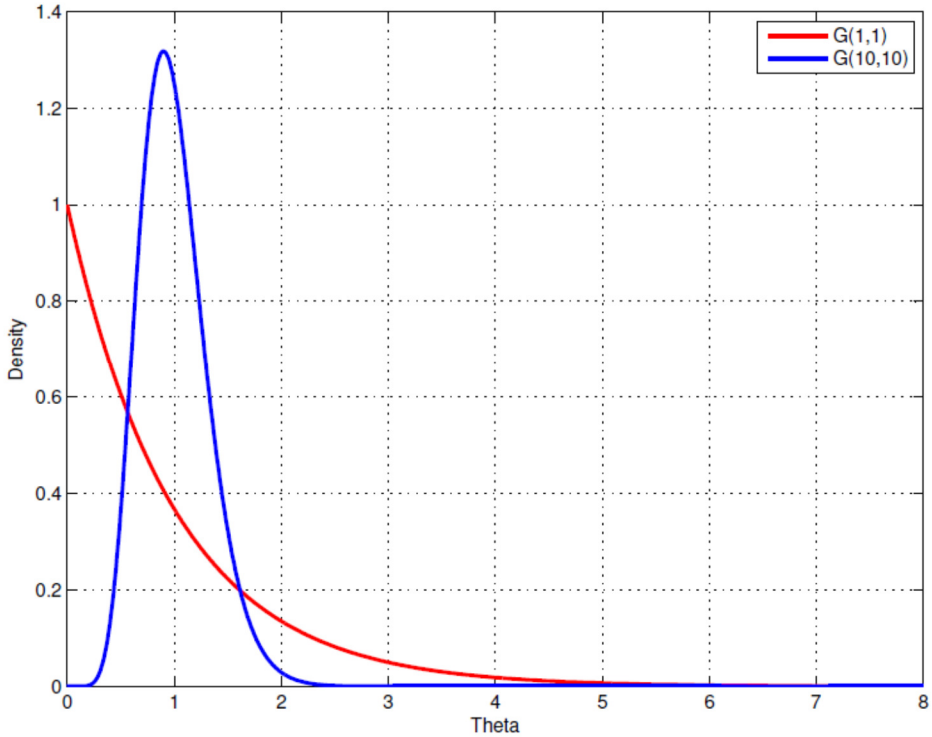

Figure 3-1 PDF of Gamma prior with parameters $\alpha = 1$, $\beta = 1$ and $\alpha = 1\,0$, $\beta = 10$.

Table 3.1 illustrates the summary statistics of closed-form and simulated posteriors under different prior distributions. Table 3.1 gives the Mean, Standard Deviation, Quantiles, Median and Estimates of posteriors' distributional parameters for θ and c. The first column of the table shows the wind speed sample size, n, and the sum of the squared randomly selected wind speed samples, T_n. It is seen that T_n increases while n increases. This always happens when the sample size is large. But this statement is not necessarily true for the small sample sizes like n = 3 and 5. We have randomly selected n samples multiple times, say one thousand times. The mean of the distribution of T_n for n = 5 is always greater than that of n = 3. We have chosen those wind speed samples which have the value of T_n nearest the mean. Table 3.1 illustrates that the mean values of the posterior distributions of the c parameter get closer to the scale parameter value of the Rayleigh distribution without uncertainty while n and T_n increase. Also the variance of the posterior distribution of the c parameter decreases when n and T_n increase.

Table 3-1 summary statistics of closed-form and simulated posteriors under different choices of prior distribution while changing n and T_n.

n, Tn	Priors	Par.		Mean	Std.	2.5%	Med.	97.5%	Est. of $G(\alpha_\theta, \beta_\theta)$
n=3 T_n=141.3158	G(1,1)	θ	Tru.	0.0285	0.0144	0.0080	0.0264	0.0655	(4.0000,142.3158)
			Sim.	0.0284	0.0165	0.0072	0.0252	0.0674	(3.0289,106.8099)
		c	Sim.	6.8288	2.6271	3.8519	6.2944	11.7590	
	G(10,10)	θ	Tru.	0.0268	0.0136	0.0075	0.0248	0.0616	(13.0000,151.3158)
			Sim.	0.1037	0.0306	0.0517	0.1000	0.1729	(11.5370,111.2580)
		c	Sim.	3.2108	0.4941	2.4049	3.1623	4.3988	
n=5 T_n=247.9977	G(1,1)	θ	Tru.	0.0240	0.0101	0.0085	0.0227	0.0485	(6.0000,248.9977)
			Sim.	0.0236	0.0106	0.0078	0.0222	0.0493	(4.7752,197.5440)
		c	Sim.	7.0756	1.8970	4.5408	6.7116	11.3228	
	G(10,10)	θ	Tru.	0.0231	0.0095	0.0081	0.0217	0.0451	(15.0000,257.9977)
			Sim.	0.0639	0.0168	0.0327	0.0628	0.1008	(13.9899,218.8347)
		c	Sim.	4.0661	0.5792	3.1497	3.9903	5.5249	
n=20 T_n=1808.2000	G(1,1)	θ	Tru.	0.0116	0.0025	0.0072	0.0114	0.0171	(21.0000,1809.2000)
			Sim.	0.0122	0.0028	0.0075	0.0120	0.0180	(19.4000,1585.2000)
		c	Sim.	9.2166	1.0710	7.4514	9.1211	11.5380	
	G(10,10)	θ	Tru.	0.0165	0.0030	0.0112	0.0163	0.0229	(30.0000,1818.2000)
			Sim.	0.0176	0.0032	0.0119	0.0174	0.0245	(29.6000,1878.8000)
		c	Sim.	7.6301	0.7115	6.3641	7.5853	9.1478	

Table 3.2 gives the summary statistics for the closed-form and simulated predictive wind speed under different choices of prior while changing n and T_n. This table illustrates that the mean value of predictive wind speed distributions are getting closer to the mean value of the Rayleigh distribution, 8.23, while increasing n and T_n under same choices of prior. From the engineering point of view, it interprets that if the sample size is small or available data provide only indirect information about the parameters of interest, the prior distribution becomes more important.

Table 3-2 Summary statistics of closed-form and simulated predictive wind speed under different choices of prior distribution while changing n and T_n.

n, Tn	Priors	Dist.	Mean	Std.	2.5%	Med.	97.5%	Est. of P(IV) (σ, γ, a)
		Ray.	8.2300	4.2205	1.4740	7.7068	17.7540	
n=3 T_n=141.3158	G(1,1)	P(IV)	5.8559	3.9765	0.9680	5.1929	14.6960	(11.9296,1/2,4)
		Pred.	5.9290	3.8250	0.9653	5.1590	15.7900	
	G(10,10)	P(IV)	3.1144	1.0251	0.6600	3.4980	7.5680	(12.3010,1/2,13)
		Pred.	2.8740	1.5930	0.4919	2.6820	6.5170	
n=5 T_n=247.9977	G(1,1)	P(IV)	6.0997	4.1362	0.9688	5.1045	13.5100	(15.7797,1/2,6)
		Pred.	6.3700	3.8950	1.0510	5.6800	15.7000	
	G(10,10)	P(IV)	3.7706	1.1473	0.9470	3.4970	11.5720	(16.0623,1/2,15)
		Pred.	3.5900	2.0220	0.6576	3.2330	8.4760	
n=20 T_n=1808.2000	G(1,1)	P(IV)	8.3764	2.1267	1.4520	7.6120	18.1940	(42.5347,1/2,21)
		Pred.	8.1780	4.4080	1.8120	7.3760	18.4900	
	G(10,10)	P(IV)	6.9871	1.4704	1.2410	7.7008	13.8160	(42.6404,1/2,30)
		Pred.	6.7240	3.6370	1.2180	6.2360	15.0800	

Figure 3.2 shows the distributions of the Rayleigh as well as the closed-forms of predictive wind speeds considering G(1,1) prior while changing n and T_n. The reference lines show the V_{CI} = 4, V_r = 14 and V_{CO} = 17 [m/s]. In the upper panel of the figure, it can be found that the PDFs are taking the form of the Rayleigh as n and T_n increase. From the lower panel of the figure, it can be seen that the probabilities of the wind speed availabilities are increasing while n increases.

Figure 3.3 shows the distributions of the Rayleigh as well as the closed-forms of predictive wind speeds considering G(10,10) prior while changing n and T_n. From the upper panel of the figure, it is also seen that the PDFs are taking the form of the Rayleigh as n and T_n increase. From the lower panel of the figure, it can be found that the probabilities of the wind power production are increasing while n increases.

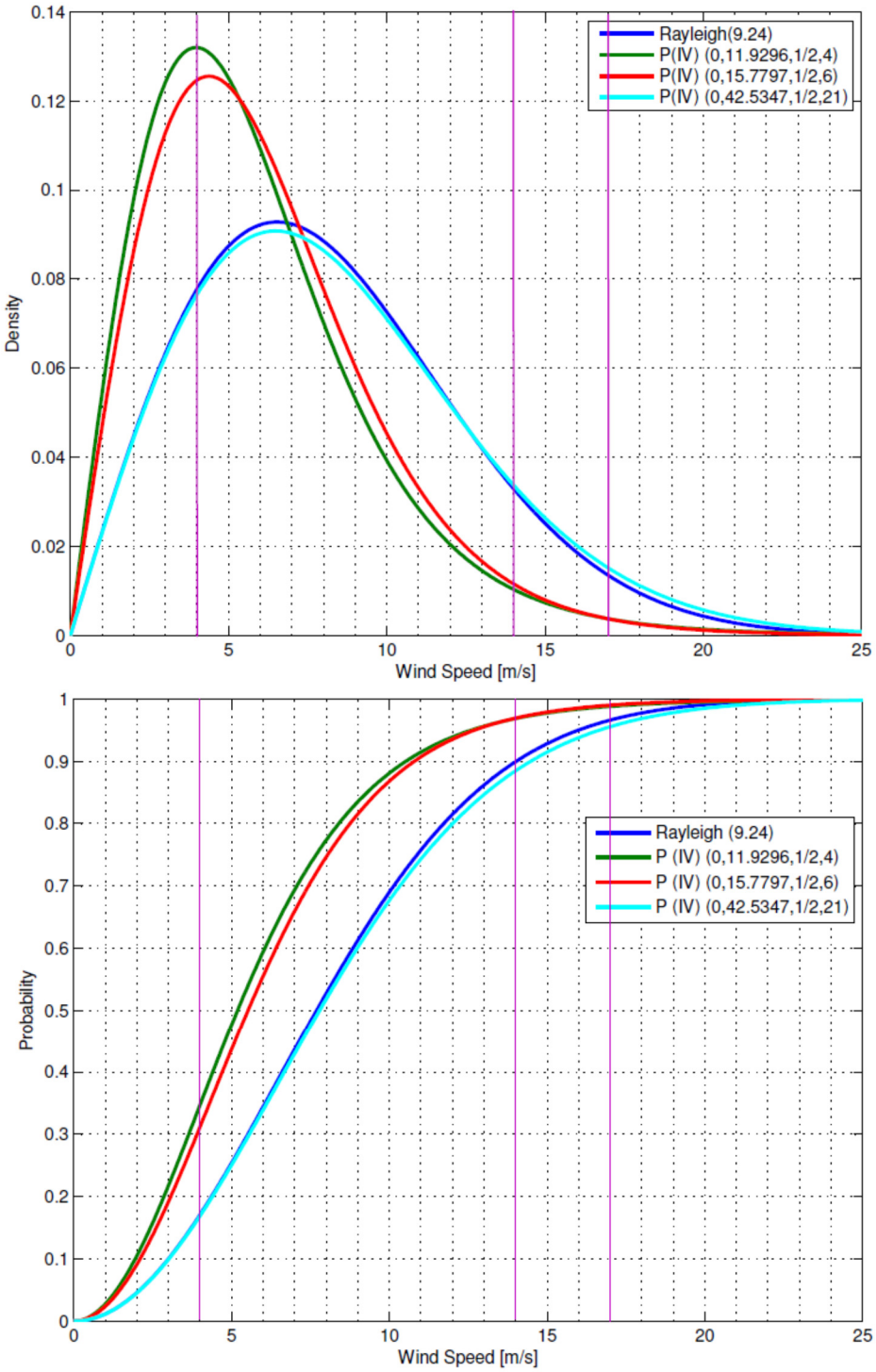

Figure 3-2 Distributions of Rayleigh and closed-from *P*(IV) of predictive wind speed considering *G*(1,1) prior.

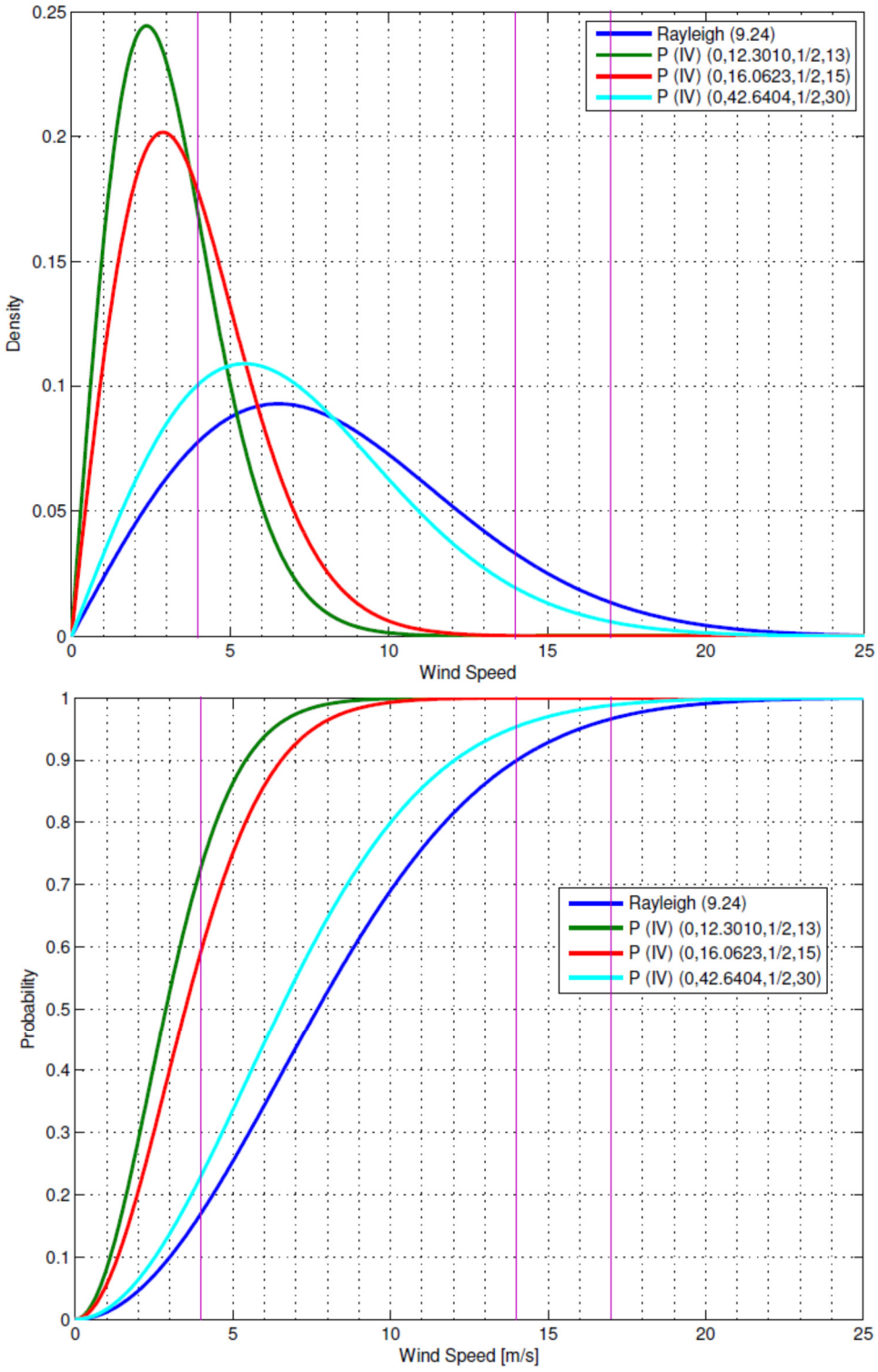

Figure 3-3 Distributions of Rayleigh and closed-from *P*(IV) of predictive wind speed considering *G*(10,10) prior.

3.2.2.1. Wind power implications

Table 3.3 shows the probabilities of wind power availability for the Rayleigh model without uncertainty as well as the predictive wind speed model under different choices of priors while changing n. More specifically the probability of available wind speed is the difference of the probabilities of the wind speed being less than or equal to V_{CO} and that less than or equal to V_{CI}. Table 3.3 also gives the probability of the maximum available wind power as the difference between the probabilities of the wind speed being less than or equal to V_{CO} and that less than or equal to V_r, in the last column. It can be seen that the availability of the wind power is 79.52% for the Rayleigh without uncertainty. The availability of the wind power are 64.11%, 67.84% and 78.67%, for the P(IV) predictive model with $G(1,1)$ prior for $n = 3$, 5 and 20 respectively. Likewise these probabilities are 27.07%, 40.55% and 75.69% for the P(IV) predictive model with $G(10,10)$ prior on the scale parameter. This table reveals that the available wind power and maximum available wind power probabilities increase to get closer to the true value of the Rayleigh model without uncertainty while n increases. Furthermore this table depicts that if the sample size is small or available data provide only indirect information about the parameters of interest, the prior distribution becomes more important.

Table 3-3 Probabilities of available wind power for Rayleigh and predictive wind speed models under different choices of prior distribution on scale parameter while changing n.

Prior	Dist.	P($W \leq V_{CI}$)	P($W \leq V_{CO}$)	P(WPA)	P($W \leq V_r$)	P(WPA)-max
	Ray.	17.08	96.61	79.52	89.93	6.68
$G(1,1)$	P(IV) (n=3)	34.70	98.81	64.11	96.86	1.94
	P(IV) (n=5)	31.18	99.02	67.84	96.93	2.09
	P(IV) (n=20)	16.88	95.54	78.67	88.47	7.08
$G(10,10)$	P(IV) (n=3)	72.93	100.00	27.07	100.00	00.00
	P(IV) (n=5)	59.45	100.00	40.55	100.00	00.00
	P(IV) (n=20)	23.11	98.80	75.69	95.36	3.44

Figure 3.4 shows the effect of changes in T_n on the CDF of the closed-form P(IV) of predictive wind speed while keeping n and prior constant, e.g. n = 3 and G(10,10). It is obvious that the probability of the available wind power increases when T_n increases.

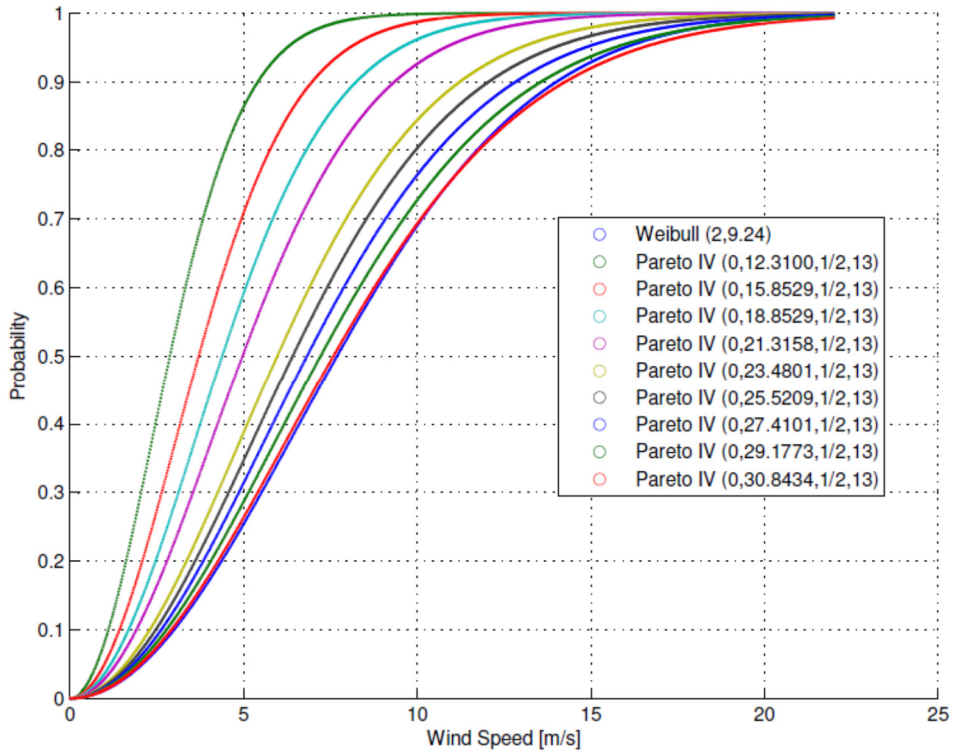

Figure 3-4 CDFs of P(IV) of predictive wind speed while changing the T_n considering n=3 and G(10,10) prior.

3.3. Example: Weibull model for data

According to the International Standard IEC 61400-12 and other international recommendations, the two-parameter Weibull PDF is the most appropriate distribution function for wind speed data, which has been used for the ‘‘construction” of the most well-fitting wind turbine model. This is essential in order to make a decision about the economic feasibility of a wind power project and, furthermore, about the selection of the most suitable site for installing a wind farm, through the estimation of the Annual Energy Production (AEP) of each wind turbine. Moreover, almost every commercial software that offers estimations of AEP is based on the two-parameter W-PDF [61]

and [67]-[69]. Weibull is a generalization of the Rayleigh distribution, and has been shown to fit wind samples better than Rayleigh due to its more flexible form (with an additional shape parameter) [64], [70]. The Weibull PDF is:

$$f\left(w \mid k, c\right) = \left(\frac{k}{c}\right)\left(\frac{w}{c}\right)^{k-1} e^{-\left(\frac{w}{c}\right)^{k}}, \tag{3-17}$$

This distribution has two parameters: k controls the shape of the distribution and c controls the scale of the distribution. The Weibull shape and scale parameters calculated using the MLE is given by:

$$\hat{k} = \left[\frac{\sum_{i=1}^{n} w_i^{\hat{k}} \ln\left(w_i\right)}{\sum_{i=1}^{n} w_i^{\hat{k}}} - \frac{\sum_{i=1}^{n} \ln\left(w_i\right)}{n}\right]^{-1} \tag{3-18}$$

where w_i is the wind speed in time step i and n is the number of data points. To evaluate (3-18) an iterative technique is used. The scale parameter is obtained by:

$$\hat{c} = \left[\frac{1}{n}\sum_{i=1}^{n} w_i^{\hat{k}}\right]^{1/k} \tag{3-19}$$

3.3.1. Shape parameter known

The model description is incomplete in Bayesian analysis until prior distributions over parameters are specified. When the shape parameter is known, the scale parameter has a conjugate Gamma prior similar to the Rayleigh model. It is assumed that the prior distribution of c is $G\left(\alpha_c, \beta_c\right)$. For simplifying the illustration we re-parameterize $\theta = \frac{1}{c^k}$ and let $\sum w_i^k = T_n$. Because w_i are not

functions of θ then, $L(\theta) \propto \theta^n e^{-\theta T_n}$. It is obvious that the only difference between a Weibull model with known shape parameter and a Rayleigh model is the value of T_n.

The survival function in the closed-form calculated from a Bayesian predictive distribution of $f(w|\theta)$ is:

$$\overline{F}(w|D) = \left(1 + \frac{w^k}{(\beta + T_n)}\right)^{-(\alpha+n)}, \tag{3-20}$$

Likewise in the Rayleigh model it can be seen that (3-20) is a Pareto Type IV distribution with scale parameter $(T_n + \beta)^{1/k}$, shape parameter k, and tail index $\alpha + n$. The PDF of the predictive wind speed in the general form with known shape parameters can be derived as follows:

$$\begin{aligned} f(w|D) &= \frac{d}{dw}\left(1 - \overline{F}(w|D)\right) \\ &= \frac{d}{dw}\left\{1 - \left(1 + \frac{w^k}{(\beta + T_n)}\right)^{-(\alpha+n)}\right\} \\ &= k(\alpha+n)(T_n+\beta)^{-1/k}\left(\frac{w}{(T_n+\beta)^{1/k}}\right)^{k-1}\left(1 + \frac{w^k}{(\beta+T_n)}\right)^{-(\alpha+n+1)} \end{aligned} . \tag{3-21}$$

The finiteness of the mean, and the existence and the finiteness of the variance depend on the tail index $\alpha + n$ and shape parameter k. The mean and variance of the expression (3-20) exist because $k < \alpha + n$.

The mean of (3-20) is:

$$E(W) = \frac{(T_n+\beta)^{1/k}.\Gamma(\alpha+n-k).\Gamma(1+k)}{\Gamma(\alpha+n)}. \tag{3-22}$$

The variance of (3-20) is:

$$Var(W) = E(W^2) - E(W)^2$$
$$= \frac{(T_n+\beta)^{2/k}.\Gamma(1+2k).\Gamma(\alpha+n-2k)}{\Gamma(\alpha+n)} - \frac{(T_n+\beta)^{2/k}.\Gamma(1+k)^2.\Gamma(\alpha+n-k)^2}{\Gamma(\alpha+n)^2}. \quad (3\text{-}23)$$

3.3.1.1. Simulation and results

In this section we simulate limited wind speed samples, say n = 3 from a Weibull model without uncertainty. We consider the wind speed model which has been developed in the second chapter with k = 3 and c = 9.15. Next we wish to consider the uncertainty about the scale parameter of the Weibull model when the shape parameter is known. Hence, we consider $G(1,1)$ and $G(10,10)$ priors on the scale parameter and keep the mean value unchanged. Table 3.4 depicts the summary statistics of the closed-form and simulated posteriors under different prior distributions while keeping n and T_n unchanged. Table 3.4 gives the Mean, Standard Deviation, Quantiles, Median and Estimates of posteriors' distributional parameters for θ and c. Table 3.4 illustrates that the mean value of the posterior distribution is closer to the scale parameter value of the Weibull distribution without uncertainty when we consider the more informative prior.

Table 3-4 Summary statistics of closed-form and simulated posteriors under different prior distributions for n=3 and T_n=2517.1.

Priors	Par.		Mean	Std.	2.5%	Med.	97.5%	Est. of $G(\alpha_\theta, \beta_\theta)$
$G(1,1)$	θ	Tru.	0.0016	0.0008	0.0004	0.0015	0.0037	(4.0000,2518.1000)
		Sim.	0.0013	0.0007	0.0004	0.0011	0.0031	(2.9990,2354.8000)
	C	Sim.	8.5499	2.2802	6.8582	9.6872	13.5721	
$G(10,10)$	θ	Tru.	0.0016	0.0008	0.0004	0.0015	0.0037	(13.0000,2527.1000)
		Sim.	0.0051	0.0015	0.0025	0.0049	0.0084	(11.5376,2276.5000)
	C	Sim.	5.8095	2.0448	4.9193	5.8875	7.3680	

Table 3.5 illustrates the summary statistics for the closed-form and simulated predictive wind speed models under different prior distributions while keeping n and T_n unchanged. Table 3.5 also illustrates that the mean value of predictive wind speed distributions get closer to the mean value of the Weibull distribution, 8.13 when we consider the more informative prior.

Table 3-5 Summary statistics of predictive inferences for the closed-form and simulation under different choices of prior distribution for n=3 and T_n=2517.1.

Priors	Dist.	Mean	Std.	2.5%	Med.	97.5%	Est. of P(IV) (σ,γ,a)
	Weib.	8.1310	2.9720	2.8790	8.1150	14.0000	
G(1,1)	P(IV)	6.6782	2.5842	2.5192	7.8103	15.6248	(0, 13.0397,1/3,4)
	Pred.	7.4290	3.1750	2.2970	7.0970	14.0500	
G(10,10)	P(IV)	5.2641	2.2944	1.7015	5.1726	9.3946	(0, 13.6210,1/3,13)
	Pred.	5.3360	2.0340	1.7040	5.2800	9.5420	

3.3.1.1.1. Wind power implications

Table 3.6 shows the probabilities of wind power availability for the Weibull model without uncertainty as well as the predictive wind speed model under different choices of priors. Table 3.6 gives the probability of available wind power as well as the probability of the maximum available wind power. It can be seen that the availability of wind power is 91.82% for the Weibull without uncertainty. The availabilities of wind power are 89.13%, 79.25% for the P(IV) predictive model with G(1,1) and G(10,10) priors respectively. This table reveals that the available wind power and the maximum available wind power probabilities are closer to the true values of Weibull when we consider the more informative prior. Furthermore Table 3.6 depicts that if the sample size is small or available data provide only indirect information about the parameters of interest, the prior distribution becomes more important.

Table 3-6 Probabilities of available wind power for predictive wind speed models under different choices of prior distribution for c while n=3 and T_n=2517.1.

Prior	Dist.	P($W \leq V_{CI}$)	P($W \leq V_{CO}$)	P(WPA)	P($W \leq V_r$)	P(WPA)-max
	Weib.	8.01	99.83	91.82	97.22	2.61
G(1,1)	P(IV)	9.55	98.68	89.13	94.76	3.92
G(10,10)	P(IV)	27.75	100.00	72.25	100.00	0

Figure 3.5 shows the distributions of the Weibull without uncertainty as well as the closed-forms of the predictive wind speeds considering different priors while keeping n unchanged. It can be seen in the upper panel of Figure 3.5 that the PDFs are taking the form of the Weibull distribution

while having the more informative prior. It can be seen in the lower panel of this figure that the probability of wind power production increases for such a statement.

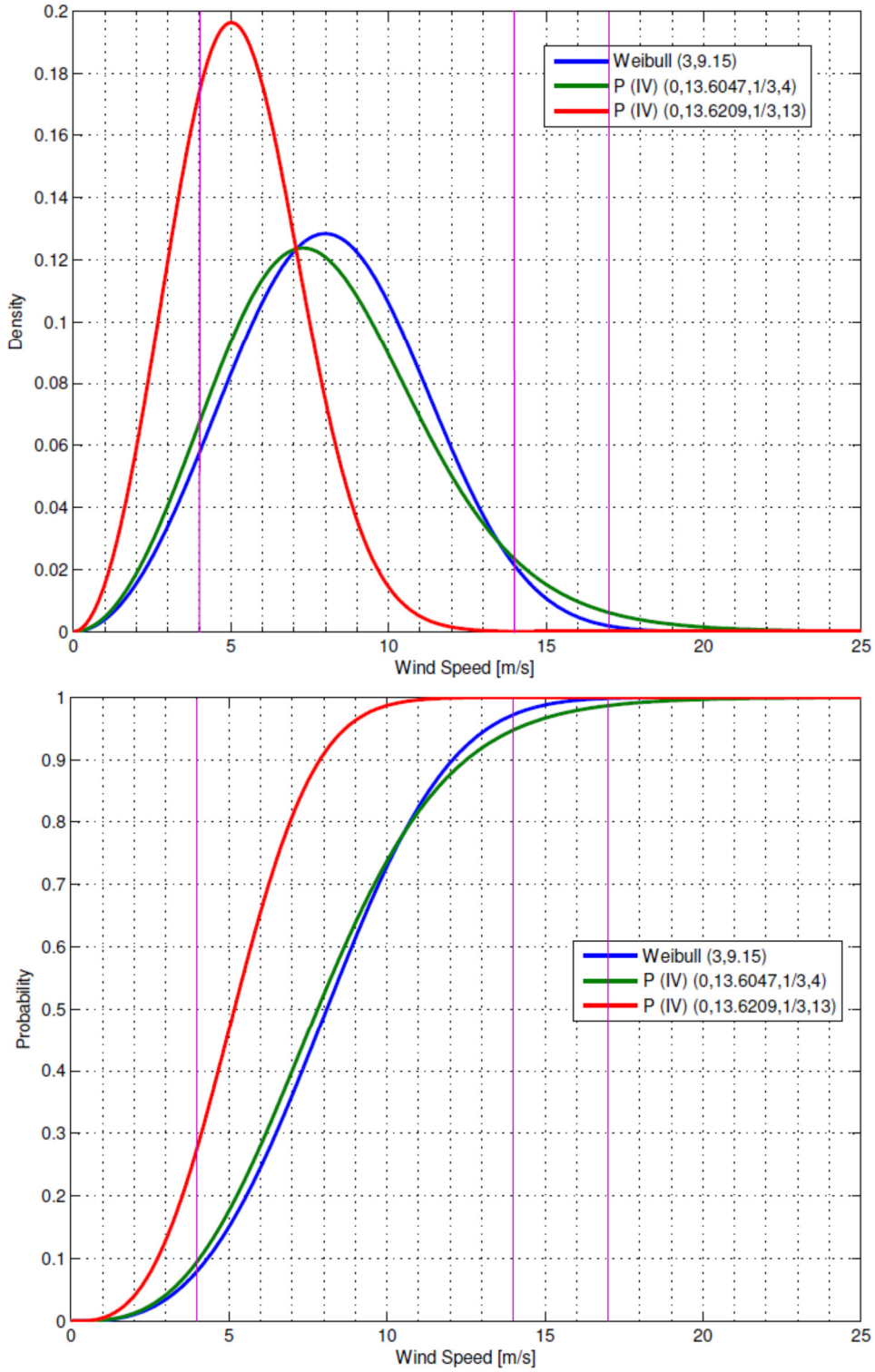

Figure 3-5 Distributions of Weibull and closed-from *P*(IV) of predictive wind speed considering n=3, T_n=2517.1, *G*(1,1) and *G*(10,10) priors.

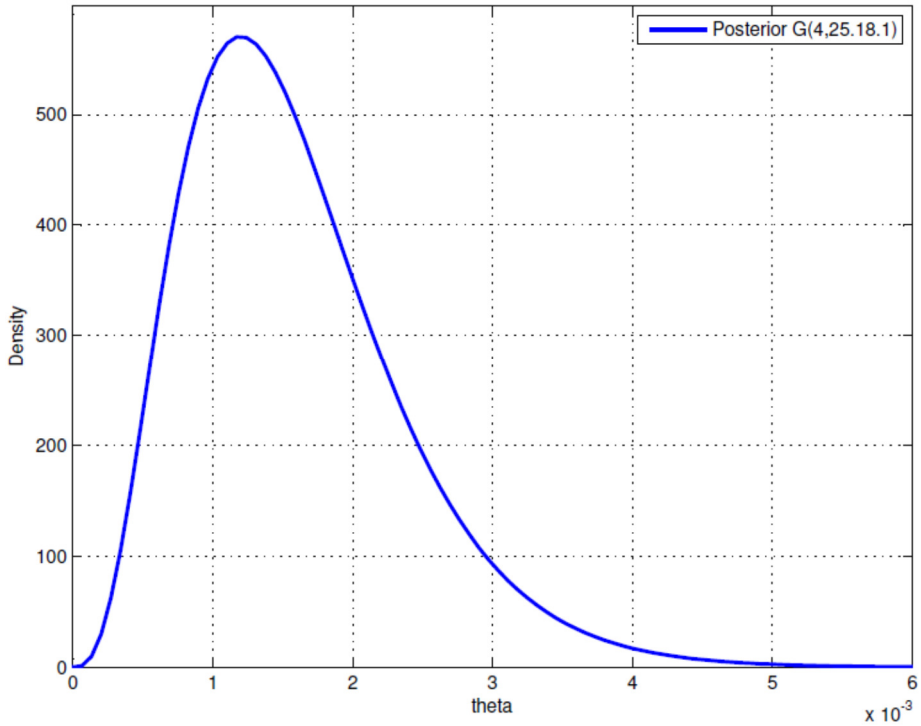

Figure 3-6 Distributions of closed-from $G(4,2518.1)$ of posterior considering $n=3$, $T_n=2517.1$ and $G(1,1)$ prior.

Figure 3.6 shows the distribution of the closed-from of the posterior, $G(4,2518.1)$, considering $n=3$, $T_n=2517.1$ and $G(1,1)$ prior. It can be seen that the posterior distribution is more informative than the prior distributions.

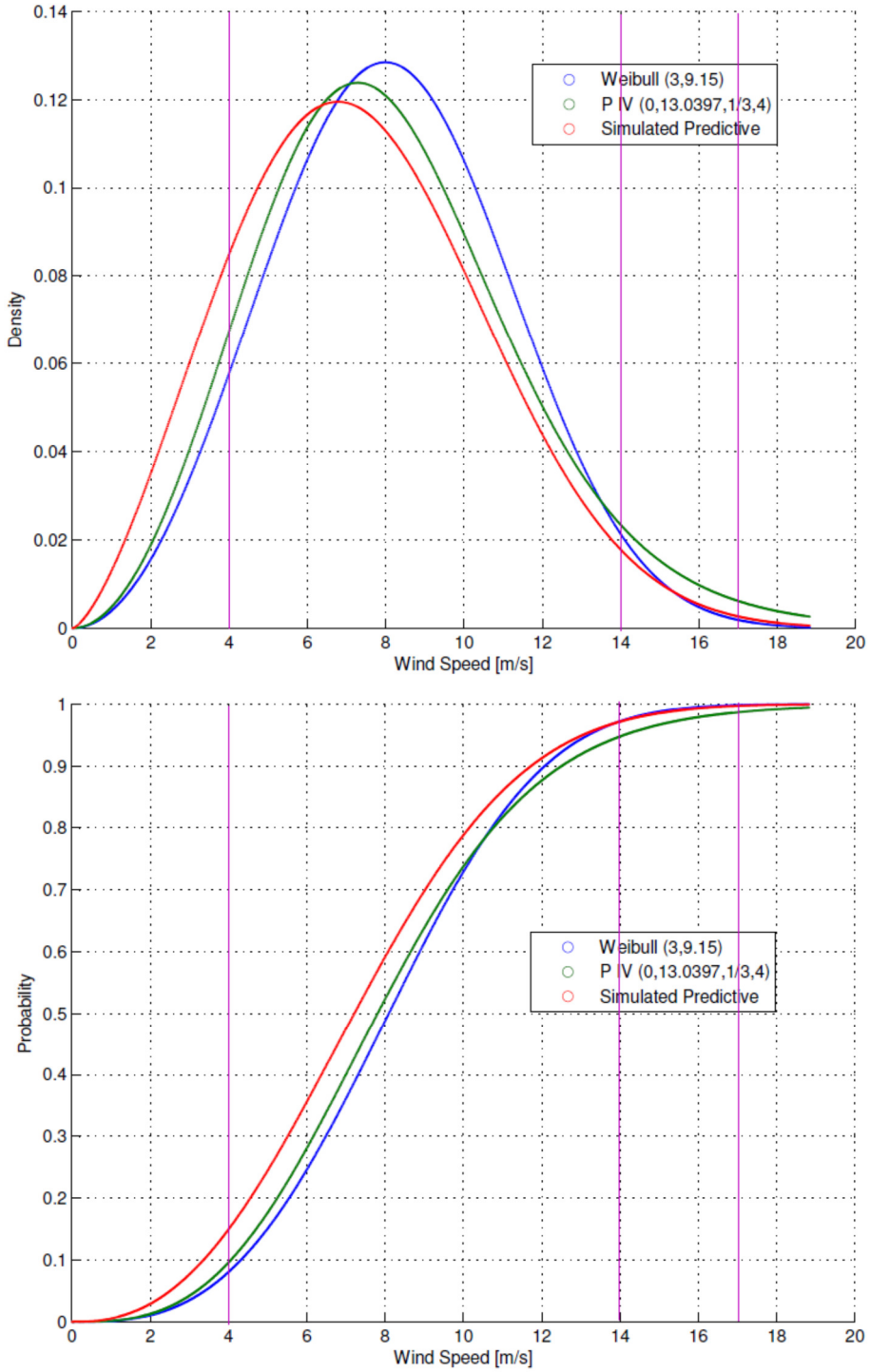

Figure 3-7 Distributions of Weibull, closed-from *P*(IV) and simulated predictive wind speed considering n=3, T_n=2517.1 and *G*(1,1) prior.

Figure 3.7 shows the distributions of the Weibull without uncertainty, the closed-from *P*(IV) and the simulated predictive wind speed. The upper panel shows a slight difference among the PDFs. Comparing the CDFs, the probabilities of wind power production for the closed-form and

simulated predictive wind speed, 89.13% and 85.13%, are less than that of the Weibull without uncertainty, 91.82%.

3.3.2. Shape parameter unknown

In this section we assume that parameters *k* and *c* both are unknown and independently distributed according to two log-concave distributions with parameters. We begin considering Gamma priors for shape and scale parameters as follows:

$$p(k) = G(\alpha_1, \beta_1), \tag{3-22}$$

$$p(c) = G(\alpha_2, \beta_2), \tag{3-23}$$

We now have a fully specified model, having a likelihood as well as a prior on both parameters. It is well known that for a Weibull distribution, while computing the Bayes estimates, the continuous conjugate joint prior distribution of the shape and scale parameters does not exist. Considering independent priors over parameters, Bayes' Theorem gives the posterior:

$$p(k, c \mid D) \propto L(k, c)\, p(k)\, p(c). \tag{3-24}$$

Expression (3-26) includes multi-dimensional integrals and therefore is not possible to be calculated analytically. Modern Bayesian analysis is typically performed by simulating the posterior distribution using MCMC. The most common application of MCMC algorithms is numerically calculating multi-dimensional integrals. It is important to realize that although the expression in (3-24) looks daunting, we can evaluate the unnormalized density of this function for any set of parameter values *k* and *c* and observations $w_1,\ldots,w_i$ (remember that we only need unnormalized densities in MCMC, which is an appealing feature of this approach).

Because the conditional distributions in (3-24) do not correspond to any analytic expression, we developed a code in WinBUGS which basically uses an MH sampler to estimate the parameters and simulates data from the posterior. The posterior distribution is bivariate as it involves two parameters *k* and *c*.

3.3.2.1. Simulations and results

This section will illustrate the results for taking the uncertainty about the model parameters, which has been developed in chapter 2, into account. For this purpose we will simulate limited wind speed samples. Then posterior and predictive inferences under different reasonable choices of prior distributions on both scale and shape parameters will be presented. We continued the simulation with 20 total wind samples and presented informative pairs Gamma priors on Weibull shape and scale parameters. First we considered the shape hyper parameters $\alpha_k = \beta_k = 1$ while varying the scale parameters according to the values $\alpha_c = \beta_c = 1$ and $\alpha_c = \beta_c = 5$. Afterwards we decreased the variance of the shape prior distribution by increasing $\alpha_k = \beta_k = 5$ while varying the scale priors according to $\alpha_c = \beta_c = 1$ and $\alpha_c = \beta_c = 5$.

Table 3-7 Summary statistics of posteriors inference under different choices of prior distribution for *k* and *c*.

p(k)	*p*(c)	Par.	Mean	Std.	2.5%	Med.	97.5%
G(1,1)	*G*(1,1)	*k*	2.6102	0.4866	1.7190	2.5780	3.6424
		λ	0.0050	0.0067	0.0002	0.0029	0.0229
		c	9.7015	0.9208	2.8203	9.6456	141.845
	G(5,5)	*k*	1.3890	0.2702	0.8936	1.3720	1.9490
		λ	0.0685	0.0459	0.0137	0.0574	0.1706
		c	8.0870	1.3651	2.4778	8.0295	120.965
G(5,5)	*G*(1,1)	*k*	2.1700	0.3932	1.4400	2.1445	3.0740
		λ	0.0114	0.0115	0.0010	0.0079	0.0422
		c	9.5461	1.4930	2.8000	9.4451	121.153
	G(5,5)	*k*	1.3210	0.2402	0.8914	1.3080	1.8310
		λ	0.0753	0.0428	0.0189	0.0664	0.1825
		c	8.0234	1.3790	2.5319	7.9102	85.7044

Table 3.7 illustrates the summary statistics of simulated posteriors under different reasonable choices of priors distributions for *k* and *c*. This table illustrates that the simulated mean values are closer to the true *k* and *c* values, 9.15 and 3, under the more informative choice of priors. Also the posteriors of the *c* parameters have heavy tails in all cases.

Table 3.8 illustrates the summary statistics of predictive inference under different choices of prior distribution for *k* and *c*. The simulated mean values are closer to the true mean value of the Weibull model, 8.13, under the more informative priors. Also the standard deviations and quantile values show that the density function of the predictive wind speed are wider under the less informative priors.

Table 3-8 Summary statistics of predictive inference under different choices of prior distribution for *k* and *c*.

p(k)	*p*(c)	Mean	Std.	2.5%	Med.	97.5%
G(1,1)	*G*(1,1)	8.6850	3.8030	2.0560	8.4420	16.7500
	G(5,5)	7.5140	6.7930	0.3338	5.9360	23.3000
G(5,5)	*G*(1,1)	8.3790	4.4640	1.3060	7.8150	19.0000
	G(5,5)	7.5577	6.4670	0.4109	6.0260	23.3400

Figure 3.8 shows the distributions of simulated predictive wind speed and posteriors considering: a) *p*(k)=*G*(1,1) and *p*(c)=*G*(1,1), b) *p*(k)=*G*(1,1) and *p*(c)=*G*(5,5), c) *p*(k)=*G*(5,5) and *p*(c)=*G*(1,1) and d) *p*(k)=*G*(5,5) and *p*(c)=*G*(5,5). This figure shows the distributions that come from MCMC in the WinBUGS software with 100,000 iterations. Figure 3.8 visualizes the results from Tables 3.6 and 3.7 perfectly.

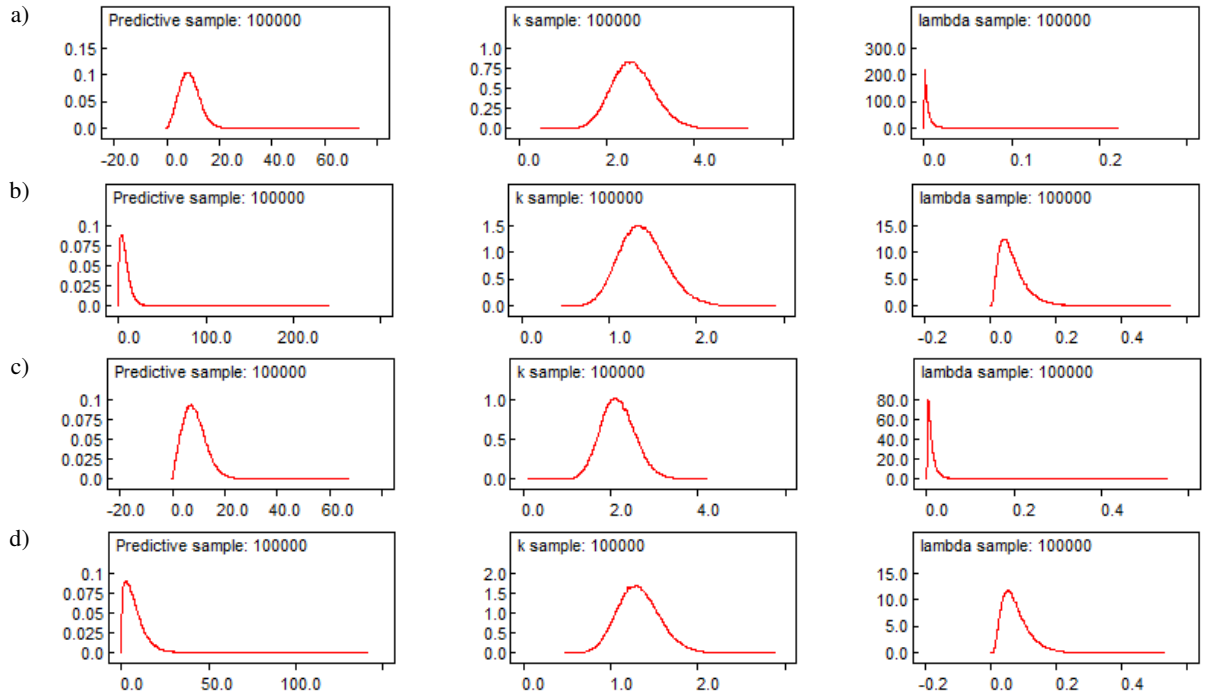

Figure 3-8 Distributions of simulated predictive wind speed and posteriors considering: a) p(k)=G(1,1) and p(c)=G(1,1), b) p(k)=G(1,1) and p(c)=G(5,5), c) p(k)=G(5,5) and p(c)=G(1,1) and d) p(k)=G(5,5) and p(c)=G(5,5).

3.3.2.1.1. Wind power implications

Table 3.9 illustrates the probabilities of available wind power for the Weibull model, the closed-form P(IV) with n = 20 and T_n = 19777 for the model with known shape parameter under different choices of parameters as well as predictive wind speed models under different choices of prior distribution for k and c while n = 20 with respect to V_{CI}, V_r and V_{CO}. The results illustrate that under a less informative choice of prior the wind power and the maximum wind power availabilities would reduce while the shape parameter is either known or unknown. From the table it can be found that the predictive wind speed distribution is more sensitive to the changes in priors of the scale parameter compared to the shape parameter while the scale parameter has more impact on the wind power and maximum availabilities.

Table 3-9 Probabilities of available wind power for Weibull, *P*(IV) with n=20 and T_n=19777 and predictive wind speed models under different choices of prior distribution for k and c while n=20.

Scale	Shape	Wind Model Dist.	P($W \leq V_{CI}$)	P($W \leq V_{CO}$)	P(WPA)	P($W \leq V_r$)	P(WPA)-max
9.15	3	*Weib.*	8.01	99.83	91.82	97.22	2.61
	3	*P*(IV)(27.0444,1/3,21)	6.55	99.05	92.50	93.47	5.58
G(1,1)	*G*(1,1)	*Pred.*	11.10	97.30	79.30	90.40	6.90
	G(5,5)	*Pred.*	14.40	94.60	72.80	87.20	7.40
	3	*P*(IV)(27.0452,1/3,25)	7.76	99.61	91.85	96.11	3.50
G(5,5)	*G*(1,1)	*Pred.*	33.30	93.20	54.20	87.50	5.70
	G(5,5)	*Pred.*	35.20	91.50	51.40	85.60	5.90

4. Chapter 4

Bayesian Applications in SEconD

The stochastic aspect of wind speed has been addressed in the literature through inclusion of the probability density function (PDF) of the wind power generation into the ED model [30]-[37], [55], [56], [58]-[60] and [69]. However, chapter 2 presents the SEconD algorithm that produces PDFs of optimal solutions in the context of Stochastic Economic Dispatching. Thus far, the effect of the uncertainty about the probability distribution of the wind speed on the distributions of optimal S-ED has not been considered. Attention to this issue is crucial in order to avoid assuming perfect knowledge of the distributional parameters which, on average, would lead to the use of unduly optimistic and insufficiently disperse distributions for the wind speed as well as optimal outputs for the planning time. This chapter reveals the effect of the uncertainty about the models' distributional parameters on the resultant outputs of the SEconD algorithm through examples. Also this chapter represents and comprehensively compares the resultant outputs of the SEconD algorithm while applied to a non-Bayesian standard wind speed model as well as a Bayesian predictive wind speed model.

4.1. Example of ED model

In this section we compare the results from a non-Bayesian and a Bayesian wind speed model applied to SEconD in order to show the effect of uncertainty in optimal outputs of the ED model. We basically carry the uncertainty about the Bayesian and non-Bayesian wind speed models through the SEconD optimizer. We use simulated predictive wind speed data points from an MCMC simulation from the Bayesian model with both shape and scale parameters unknown as well as simulated wind speed data points from an MC simulation from a non-Bayesian model

which is developed in chapter 2, for inputs to the SEconD model. Hence we produce data for estimating the probability distributions of optimal fossil fuel generation outputs, transmission loss, and total cost of power generation. The MCMC and MC samples can be used to estimate distributional functions such as PDFs, cumulative probability distributions, reliability functions (survival functions), and summary measures such as the mean, standard deviation, median, percentiles, prediction intervals, correlation coefficients, and scatter plots. Likewise large samples of simulated data allow statistical analyses of the outputs such as distributional measures and confidence intervals for each output variable and comparisons of distributions of all output variables.

In this illustration, we use a wind farm where $V_{CI} = 4$ m/s, $V_r = 14$ m/s, $V_{CO} = 17$ m/s, and $W_r = 30$ MW. Table 4.1 shows the operation ranges, cost coefficients, and the transmission loss matrices used in this example. The unit of P_i is MW, and the units of a_i, b_i and c_i are respectively \$/MW^2 h, \$/MWh, and \$/h. From chapter 2, the forecast value combined with the error distribution via (2-5) gives the Weibull distribution shown in Figure 4.1.

Table 4-1 Operation ranges and power cost coefficients.

Generator	Range		Power Cost Coefficients		
	P_{min}	P_{max}	a_i	b_i	c_i
1	54	110	0.00633	11.669	213.1
2	50	120	0.00889	10.333	200.0
3	47	100	0.00741	10.833	240.0

$$\left[B_{ij}\right] = \begin{bmatrix} 0.05370 & 0.00453 & 0.00207 & 0.04507 \\ 0.00453 & 0.06760 & 0.00953 & 0.00507 \\ 0.00207 & 0.00953 & 0.05210 & 0.00901 \\ 0.04507 & 0.00507 & 0.00901 & 0.2940 \end{bmatrix}$$

$$\left[B'_{oj}\right] = \begin{bmatrix} 0.00330 & 0.07660 & 0.00342 & 0.01890 \end{bmatrix}$$

$$\left[B_0\right] = \left[0.040357\right]$$

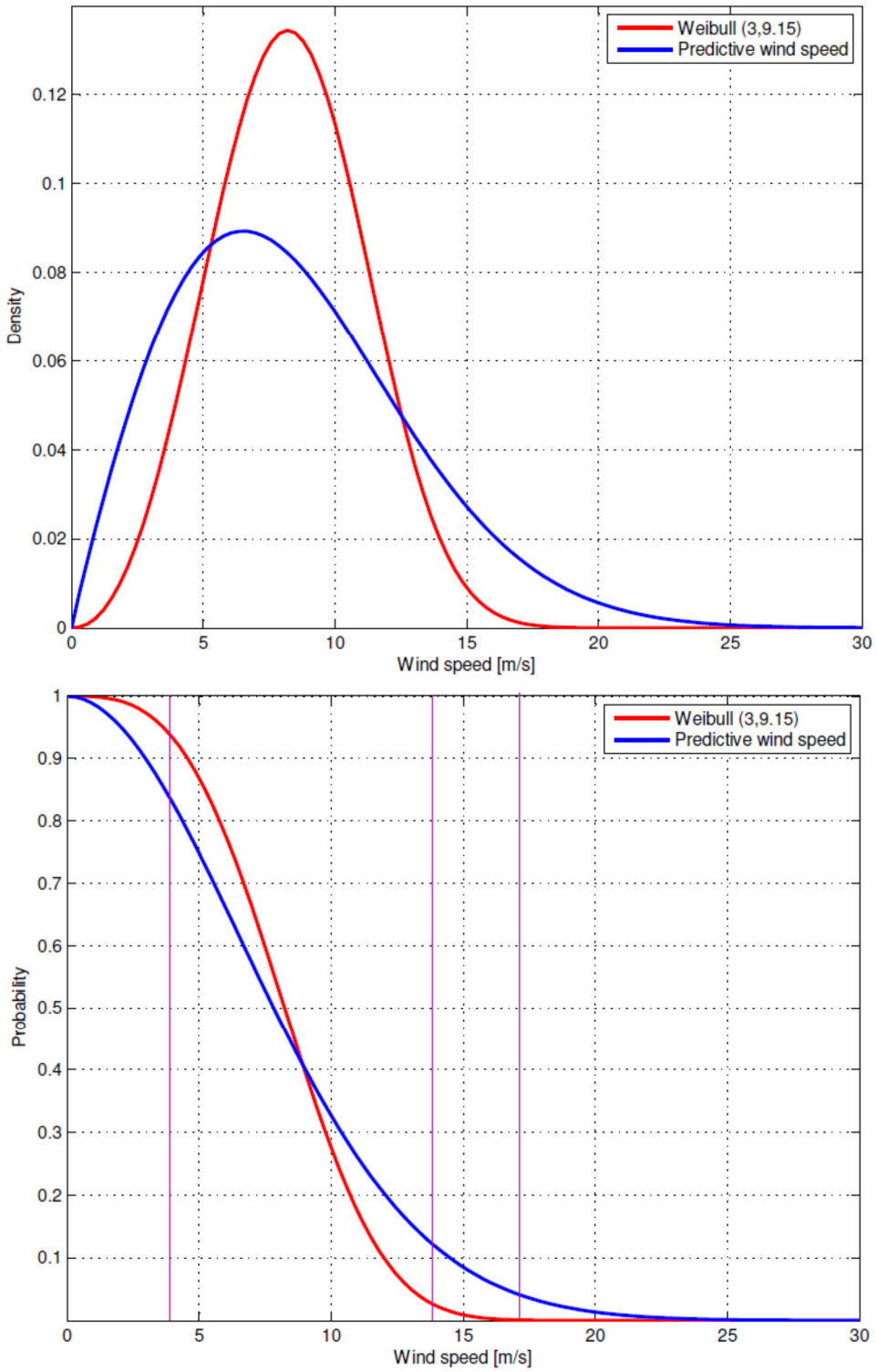

Figure 4-1 Distributions of the wind speeds from the Weibull model and simulated predictive.

Also the distribution of simulated predictive wind speed from the case with $p(k) = G(5,5)$ and $p(c) = G(1,1)$ priors is shown in Figure 4.1. It can be seen in the lower panel of the figure that the

probability of wind power generation for the Weibull model is 95% which is higher than that of the predictive model that is 78%. This 17% difference illustrates that the uncertainty about the predictive model causes less power production from the wind farm.

4.1.1. Probability distribution of optimal outputs

Table 4.2 reveals the summary statistics for the inputs (simulated wind speed data and wind power generation) and the optimal outputs of the three generators, the operation cost, and the transmission loss. Note that the ranges of the optimal outputs of the generators (between the minimum and maximum) are subintervals of their operation ranges. Comparing the means or the medians provides the following common information for both non-Bayesian and Bayesian scenarios: on average, the optimal output of generator 1 is higher than the optimal output of generator 2, which in turn is higher than the optimal output of generator 3. The upper panel of Figure 4.2 shows the PDF plots of the distributions of the optimal output powers of the three fossil fuel generators under the two scenarios.

Table 4-2 Summary statistics of non-Bayesian and Bayesian scenarios.

		Min	Max	Median	Mean	Std.
NB	*WS*	0.00	17.95	8.18	8.26	2.84
	P_w	0.00	30.00	12.53	12.88	8.02
	P_1	73.23	82.02	78.68	78.25	2.10
	P_2	66.54	75.98	74.34	73.50	1.89
	P_3	48.11	74.98	53.62	53.35	4.21
	P_L	9.81	10.31	10.17	10.15	0.09
	OC	2630.00	3112.70	2946.40	2933.79	99.64
B	*WS*	0.00	27.23	7.81	8.38	4.46
	P_w	0.00	30.00	10.01	11.48	9.84
	P_1	71.01	83.82	79.69	79.27	2.75
	P_2	65.70	77.47	74.41	72.00	2.34
	P_3	40.20	71.71	57.12	57.73	5.84
	P_L	9.78	10.35	10.19	10.11	0.12
	OC	2480.29	3195.08	2969.32	2953.01	126.30

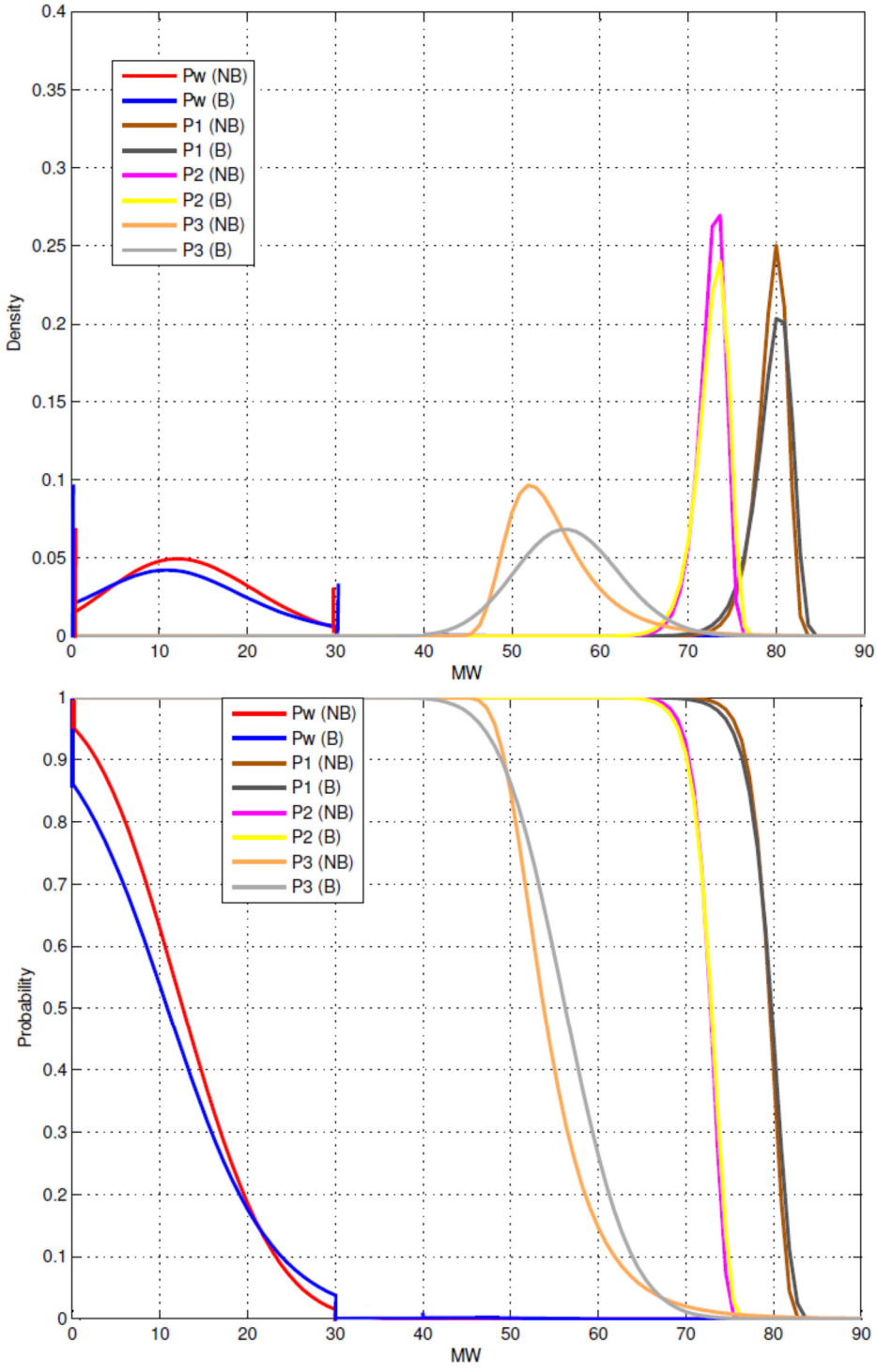

Figure 4-2 Distributions of the wind power and optimal output power generation for non-Bayesian and Bayesian scenarios.

The PDF plots indicate higher uncertainty about the power generation for the Bayesian scenario vs the non-Bayesian especially in the third generator. Consequently, the system operator may consider setting aside more reserve for this generator while considering the Bayesian scenario. The

most probable values of power outputs P_1, P_2, and P_3 are around the modes of the distribution 80, 73 and 52 MW for non-Bayesian and 80.1, 73.1 and 56 MW for Bayesian scenarios respectively. The reliability plots are shown in the lower panel of Figure 4.3. These reliability plots show the stochastic dominance of P_1 over P_2 and P_3 as well as P_2 and P_3 in both scenarios. For outputs less than 48 and more than 78 MW, P_3 in non-Bayesian dominates P_3 Bayesian and vice versa for outputs between 48 and 74 MW.

Figure 4.3 shows the PDFs and reliability functions of the distributions of the system transmission loss for both scenarios. The plots of PDFs indicate that the most likely MW values of the transmission losses are almost 10.2 and 10.15 MW for both non-Bayesian and Bayesian scenarios, respectively. However, the plots of the reliability functions provide a more complete comparison. These plots show the stochastic dominance of the transmission loss of the non-Bayesian scenario over the Bayesian before 10.15 MW. Figure 4.4 shows the PDFs and reliability functions of the distributions of the minimal cost of power production for both scenarios. The PDFs for the two scenarios cross twice, and thus do not provide a definite comparison for the two operation costs. However, the plots of the reliability functions provide a definite comparison: The optimal operation cost of Bayesian is stochastically larger than that of the non-Bayesian scenario from 2820 $/h and on. That is, for any given value x $/h greater than 2820 $/h, the probability that the optimal operation cost exceeds x is higher for the Bayesian scenario than for the non-Bayesian scenario. For instance, the probability that the optimal operation cost exceeds 3000 $/h is about 30% and 40% for non-Bayesian and Bayesian scenarios, respectively.

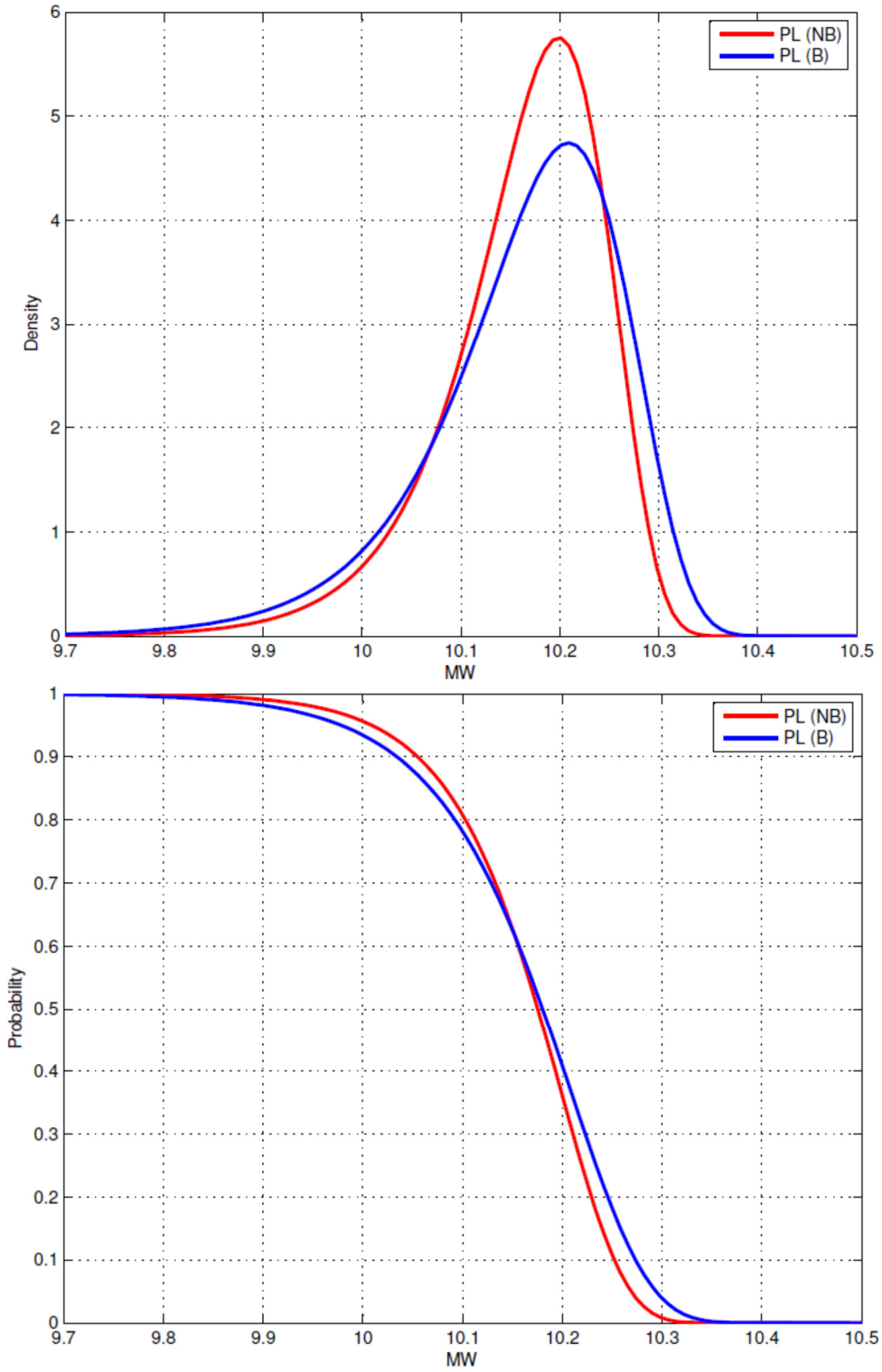

Figure 4-3 Distributions of the transmission losses for non-Bayesian and Bayesian scenarios.

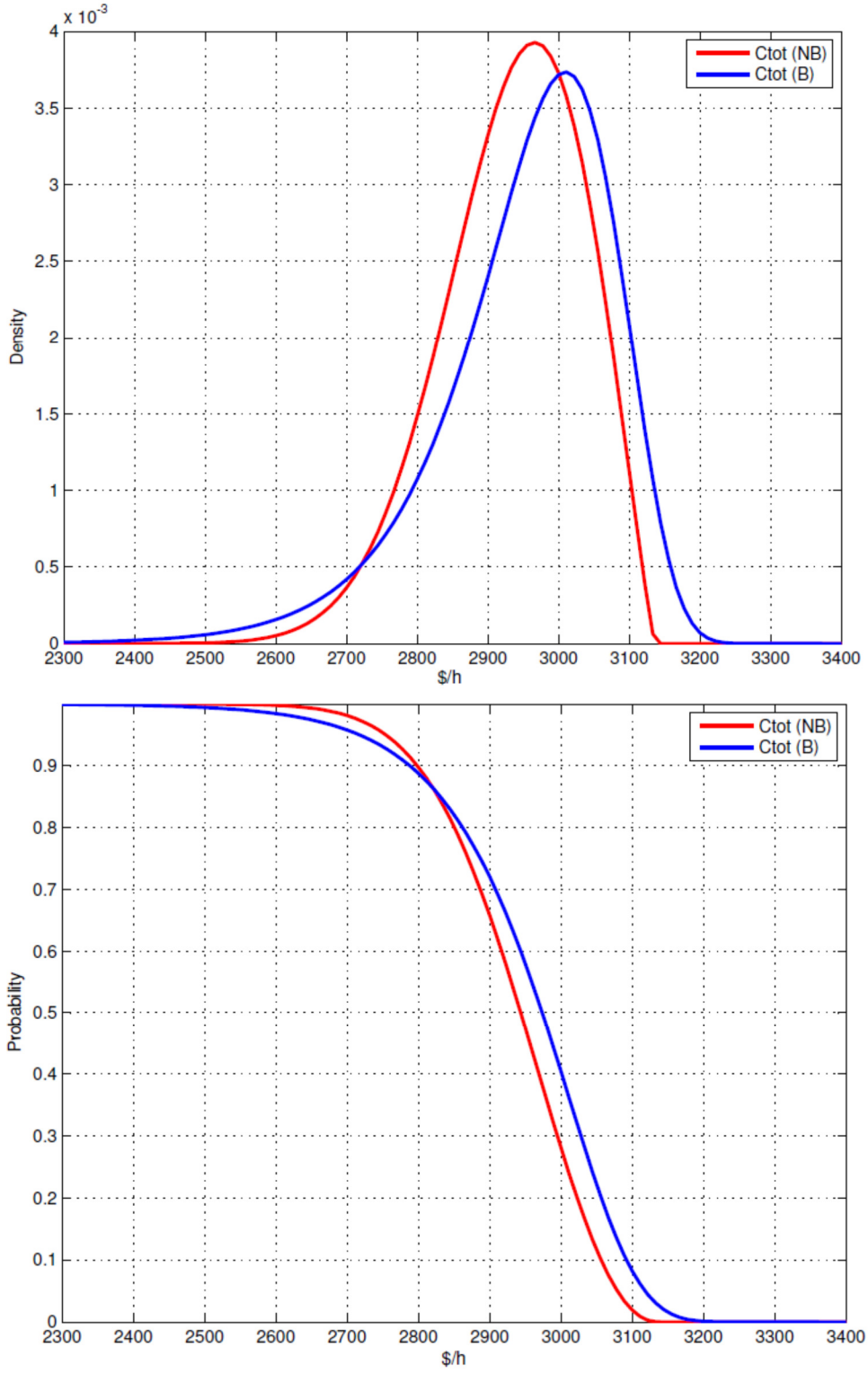

Figure 4-4 Distributions of the total costs for non-Bayesian and Bayesian scenarios.

Table 4.3 gives the correlation coefficients between the optimal outputs of the three generators, transmission loss and operation cost along with the correlation with the wind power generation

and the wind speed for non-Bayesian and Bayesian scenarios. The correlations among the variables are high for both scenarios. The correlations between wind power and the optimal fossil fuel generators' outputs are negative, most strongly with P_3. It should be noted that the correlation between wind speed and wind power is strong, but not perfect due to the cut-in and cut-out wind speeds for the non-Bayesian scenario. But for the Bayesian scenario the correlation between wind speed and wind power decreased dramatically due to the model uncertainty as well as the cut-in and cut-out wind speeds. Also there is a slight increase in the correlation of almost all variables between the non-Bayesian and Bayesian scenarios. This could happen due to model uncertainty again. While the probability of wind power generation decreased the correlation among the variables of the Bayesian scenario slightly increased.

Table 4-3 Correlation coefficients of non-Bayesian and Bayesian scenarios.

		WS	P_w	P_1	P_2	P_3	P_L	*OC*
	WS	1						
	P_w	0.97	1					
	P_1	-0.90	-0.93	1				
NB	P_2	-0.90	-0.93	0.96	1			
	P_3	-0.93	-0.96	0.80	0.80	1		
	P_L	-0.88	-0.90	0.99	0.98	0.76	1	
	OC	-0.97	-1.00	0.93	0.93	0.93	0.90	1
	WS	1						
	P_w	0.64	1					
	P_1	-0.64	-0.94	1				
B	P_2	-0.65	-0.95	0.98	1			
	P_3	-0.59	-0.96	0.81	0.82	1		
	P_L	-0.63	-0.92	0.99	0.98	0.79	1	
	OC	-0.64	-1.00	0.94	0.94	0.96	0.92	1

5. Chapter 5

Conclusion and Future Work

The quality of wind speed modeling depends on the suitability of the chosen probability models to describe the wind speed frequency distribution. The models most commonly used in recent literature to describe wind speed distribution are Weibull and Gamma, both of which belong to the Generalized Gamma (GG) family. These models view the distributions of the wind speed over a wind farm as being homogenous. However, a wind farm has multiple turbines installed in different locations, each of which may have its own distribution model. The main aim of this dissertation was to develop a wind speed model that can aggregate the non-homogenous distributions into a single continuous distribution. The secondary aim was to develop a stochastic economic dispatching algorithm that can consider the entire wind speed PDF rather than just a single randomly drawn point. This algorithm is called SEconD.

In order to test and show the capability of SEconD, we also proposed a wind speed model which uses a wind speed forecast, and an error distribution for capturing the uncertainty about the wind speed for the planning target time. It was shown that SEconD utilizes Monte Carlo simulated samples from the entire proposed wind speed spectrum as inputs to determine optimal dispatch of the corresponding fossil power plants. The SEconD algorithm is capable of incorporating any wind speed model and any ED model proposed in the literature. Some examples of the usefulness of the SEconD algorithm for the planning purposes have been illustrated. We also showed that the availability of the probability distributions of the optimal outputs enables a system operator to compare fossil fuel power generators, the transmission losses, and the operations costs. The reliability functions provided comparisons of the entire distributions of the optimal outputs of

fossil fuel power generators in terms of the stochastic dominance, and the scatter plots provided information about their relationships. Examples also illustrated that the combination of the reliability and scatter plots provides a powerful tool for the system operator to perform useful analyses. Furthermore, the distributions of the optimal outputs under two fuel price scenarios were compared. The reliability function provided definite ordering of the cost distributions under the two scenarios, which could not be detected through the PDFs.

Next we developed Bayesian predictive models to capture the uncertainty about the wind speed parameters. The closed-forms of posteriors and predictive wind speeds consider the conjugate priors for Rayleigh and Weibull models have been derived. The closed-form for the Weibull model exists when the shape is known. The MCMC simulation gave the posterior and predictive wind speed distributions for the Weibull model with both parameters unknown and therefore no longer exists in the closed-form. Comparing posterior and predictive inferences under different reasonable choices of prior distribution in sensitivity analysis, (and, for that matter, different reasonable choices of probability models for data) showed that if the sample size is small, or if the available data provide only indirect information about the parameters of interest, the prior distribution becomes more important.

The Bayesian predictive wind speed distribution was applied to the SEconD algorithm and the resultant output distributions have been comprehensively compared to non-Bayesian distributions of optimal outputs. Likewise reliability functions provided definite ordering of the cost distributions under the two scenarios, which could not be detected through the PDFs. The analyses showed the effect of uncertainty on distributions of optimal outputs. Examples also revealed that

the developed Bayesian models provide prudent information for the system operator to perform useful analyses while there is limited, incomplete or partial data available.

In this dissertation, there are quite a number of things have not yet been considered and we are pleased to offer as future work:

1- We have considered the uncertainty of the parameters of the wind speed model through the prior distributions. We used Gamma prior with known hyper-parameters. In fact there is not always perfect knowledge of prior distributions available. Therefore different levels of uncertainty can be achieved while taking into account the hyper-parameters of the prior distribution as being random.

2- We have modeled a single wind farm in a wind-penetrated power system and presented the resultant S-ED. The next goal is to model multiple dependent wind farms. In this case, a copula, which is a multivariate probability distribution, is the solution to model wind speed. Copulas are used to describe the dependence between random variables. There are many parametric copula families available, which usually have parameters that control the strength of dependence.

3- Solar power is another source of energy that similarly has randomness. Likewise we wish to develop models with different levels of uncertainty for a hybrid power system incorporating multiple solar and wind farms.

4- Technically, the SEconD algorithm is capable of incorporating any wind speed models. We wish to enhance the capability of the SEconD algorithm to enable it to incorporate other renewable sources of energy as well.

REFERENCES

[1] E. C. Morgan, M. Lackner, R. M. Vogel and L.G. Baise, "Probability distributions for offshore wind speeds", Energy Conversion and Management, vol. 52, pp 15–26, July 2011.

[2] B. G. Brown, R. W. Katz and A. H. Murphy, "Time Series Models to Simulate and Forecast Wind Speed and Wind Power," Journal of Climate and Applied Meteorology, vol. 23, pp. 1184–1195, 1984.

[3] R. Kretzschmar, P. Eckert, D. Cattani and F. Eggimann, "Neural Network Classifiers for Local Wind Prediction," Journal of Applied Meteorology, vol. 43, pp. 727–738, 2004.

[4] T. Gneiting, K. Larson, K. Westrick, M. G. Genton and E. Aldrich, "Calibrated Probabilistic Forecasting at the Stateline Wind Energy Center: The Regime-Switching Space-Time Method," Journal of the American Statistical Association, vol. 101, pp. 968–979, 2006.

[5] M. Genton and A. Hering, "Blowing in the Wind," Significance, vol. 4, pp. 11–14, 2007.

[6] G. Giebel, R. Brownsword and G. Kariniotakis, "The State-of-the- Art in Short-Term Prediction of Wind Power," Deliverable Report D1.1, Project Anemos. Available at http://anemos.cma.fr/download/ANEMOS_D1.1_StateOfTheArt_v1.1.pdf, 2003.

[7] T. N. Palmer, "The Economic Value of Ensemble Forecasts as a Tool for Risk Assessment: From Days to Decades," Quarterly Journal of the Royal Meteorological Society, vol. 128, pp. 747–774, 2002.

[8] T. Gneiting and A. E. Raftery, "Weather Forecasting With Ensemble Methods," Science, vol. 310, pp. 248–249, 2005.

[9] E. P. Grimit and C. F. Mass, "Initial Results of a Mesoscale Short- Range Ensemble Forecasting System Over the Pacific Northwest," Weather and Forecasting, vol. 17, pp. 192–205, 2002.

[10] R. Buizza, P. L. Houtekamer, Z. Toth, G. Pellerin, M. Wei and Y. Zhu, "A Comparison of the ECMWF, MSC and NCEP Global Ensemble Prediction Systems," Monthly Weather Review, vol. 133, pp. 1076–1097, 2005.

[11] Ravindra Kollu, Srinivasa Rao Rayapudi, SVL Narasimham and Krishna Mohan Pakkurthi, "Mixture probability distribution functions to model wind speed distributions," International Journal of Energy and Environmental Engineering, Springer, vol. 3, no. 27, 2012.

[12] M. R. Islam, R. Saidur and N. A. Rahim, "Assessment of wind energy potentiality at kudat and Labuan, Malaysia using weibull distribution function," Energy, vol. 36, no. 2, pp. 985–992, 2011.

[13] A. N. Celik, "Energy output estimation for small-scale wind power generators using Weibull-representative wind data," J Wind Eng Ind Aerodyn, vol. 91, no. 5, pp. 693–707, 2003.

[14] A. N. Celik, "A statistical analysis of wind power density based on the Weibull and Rayleigh models at the southern region of Turkey," Renew Energy, vol. 29, pp. 593–604, 2003.

[15] R. E. Luna and H. W. Church, "Estimation of long term concentrations using a 'universal' wind speed distribution," J Appl Meteorol, vol. 13, no. 8, pp. 910–916, 1974.

[16] A. Garcia, J. L. Torres, E. Prieto, and A. De Francisco, "Fitting wind speed distributions: a case study," Solar Energy, vol. 62, no. 2, pp. 139–144, 1998.

[17] P. Kiss and I. M. Jánosi, "Comprehensive empirical analysis of ERA-40 surface wind speed distribution over Europe" Energy Conversion and Management, vol. 49, pp. 2142–2151, 2008.

[18] P. Embrechts, C. Klüppelberg and T. Mikosch, Modelling Extremal Events for Insurance and Finance, Springer, New York, 1997.

[19] S. Kotz, and S. Nadarajah, Extreme Value Distributions: Theory and Applications, World Scientific Publishing Company, Singapore, 2001.

[20] A. Ying, and M. D. Pandey, "The r largest order statistics model for extreme wind speed estimation," J Wind Eng Ind Aerodyn, vol. 95, no. 3, pp. 165–182, 2007.

[21] P. Hyun Woo and S. Hoon, "Parameter estimation of the generalized extreme value distribution for structural health monitoring," Probabilistic Engineering Mechanics, vol. 21, no. 4, pp. 366–376, 2006.

[22] J. A. Carta and P. Ramı'rez, "Analysis of two-component mixture Weibull statistics for estimation of wind speed distributions," Renew Energy, vol. 32, no. 3, pp. 518–531, 2007.

[23] P. Kiss, I. M. Janosi, "Comprehensive empirical analysis of ERA-40 surface wind speed distribution over Europe," Energy Convers Manage, vol. 49, no. 8, pp. 2142-2151, 2008, doi:10.1016/j.enconman.2008.02.003.

[24] L. Auwera, F. Meyer and L. Malet, "The use of the Weibull three-parameter model for estimating mean wind power densities," J Appl Meteorol, vol. 19, pp. 819-825, 1980.

[25] E. W. Stacy and G. A. Mihram, "Parameter estimation for a generalized Gamma Distribution," Technometrics, vol. 7, no. 3, pp. 349-358, 1965.

[26] O. A. Jaramillo and M. A. Borja, "Wind speed analysis in La Ventosa, Mexico: a bimodal probability distribution case," Renew Energy, vol. 29, no. 10, pp. 613–630, 2004.

[27] S. Akpinar, E. K. Akpinar, "Estimation of wind energy potential using finite mixture distribution models," Energy Conversion Management, vol. 50, no. 4, pp. 877–884, 2009.

[28] C. Tian Pau, "Estimation of wind energy potential using different probability density functions," Applied Energy, vol. 88, no.5, pp. 1848–1856, 2011.

[29] S. A. Akdağ, H. S. Bagiorgas and G. Mihalakakou, "Use of two-component Weibull mixtures in the analysis of wind speed in the Eastern Mediterranean," Applied Energy, vol. 87, no. 8, pp. 2566–2573, 2010.

[30] B. H. Chowdhury and S. Rahman, "A review of recent advances in economic dispatch," IEEE Transactions on Power Systems, vol. 5, no.4, pp. 1248-1259, Nov. 1990.

[31] J. J. Grainger and W.D Stevenson, Jr., "Power system analysis, " McGraw-Hill, Inc.: New York, 1994.

[32] J. Hetzer, D. C. Yu, and K. Bhattarai, "An economic dispatch model incorporating wind power," IEEE Transactions on Energy Conversion, vol. 23, no. 2, pp. 603-611, June 2008.

[33] X. Liu, W. Xu. "Minimum emission dispatch constrained by stochastic wind power availability and cost," IEEE Transactions on Power Systems, vol. 25, No. 3, August 2010.

[34] R. A. Jabr and B. C. Pal, Intermittent wind generation in optimal power flow dispatching, IET Generation Transmission Distribution vol. 3, no. 1, pp. 66–74, 2009.

[35] D. Villanueva, A. Feijoo, and J. L. Pazos, Simulation of correlated wind speed data for economic dispatch evaluation, IEEE Trans. Sustain. Energy, vol. 3, no. 1, pp. 142-149, January 2012.

[36] N.Yorino, H. M. Hafiz, Y. Sasaki, and Y. Zoka, High-speed real-time dynamic economic load dispatch, IEEE Trans. Power Syst., vol. 27, no. 2, pp. 621-630, May 2012.

[37] P. Zhang and S. T. Lee, Probabilistic load flow computation using the method of combined cumulants and Gram-Charlier expansion, IEEE Trans. Power Syst., vol. 19, no. 1, pp. 676–682, Feb. 2004.

[38] A. Schellenberg, W. Rosehart, and J. Aguado, Cumulant-based probabilistic optimal power flow (P-OPF) with Gaussian and gamma distributions, IEEE Trans. Power Syst., vol. 20, no. 2, pp. 773–781, May 2005.

[39] R. N. Allan, A. M. L. Da Silva, and R. C. Burchett, Evaluation methods and accuracy in probabilistic load flow solutions, IEEE Trans. Power App. Syst., vol. PAS-100, no. 5, pp. 2539–2546, May 1981.

[40] G. J. Anders, Probabilistic Concepts in Electric Power Systems. Hoboken, NJ: Wiley, 1990.

[41] A. Schellenberg, W. Rosehart, and J. Aguado, Introduction to cumulant-based probabilistic optimal power flow (P-OPF), IEEE Trans. Power Syst., vol. 20, no. 2, pp. 1184–1186, May 2005.

[42] C. Monteiro, R. Bessa, V. Miranda, A. Botterud, J. Wang, and G. Conzelmann, Wind power forecasting: State-of-the-art 2009, Argonne Nat. Lab., US Department of Energy, 2009.

[43] J. S. Rohatgi, and V. Nelson, "Wind Characteristics: An analysis for the generation of Wind Power," Alternative Energy Institute Canyon, TX, 1994.

[44] J. F. Manwell, J. G. McGowan and A. L. Rogers, "Wind energy explained, theory, design and application," 2nd Edition, John Wiley & Sons Ltd., West Sussex, UK. 2009.

[45] M. Shaked and J. G. Shanthikumar, Stochastic orders, Springer, Netherlands, 2007.

[46] I. G. Damousis, M. C. Alexiadis, J. B. Theocharis, and P. S. Dokopoulos, "A fuzzy model for wind speed prediction and power generation in wind parks using spatial correlation", IEEE Transaction on Energy Conversion, vol. 19, no. 2, pp. 352-361, June 2004.

[47] B. G. Brown, R. W. Katz, and A. H. Murphy, "Time series models to simulate and forecast wind speed and wind power," Journal of Climate and Applied Methodology, vol. 23, pp. 1184-1195, Aug. 1984.

[48] S. Li, D. C. Wunsch, E. A. O'Hair, and M. G. Giesselmann, "Using neural networks to estimate wind turbine power generation," IEEE Transaction on Energy Conversion, vol. 16, no. 3, pp. 276-282, Sept. 2001.

[49] A.S. Hering, and M.G. Genton, Powering up with space-time wind forecasting, J. Am. Stat. Assoc. (JASA), vol. 105, pp. 92- 104, 20 10.

[50] X. Zhu and M.G. Genton, Short-term wind speed forecasting for power system operations, IAMCS Tech. Rep., 2011.

[51] Wind Turbine Generator Systems. Part 1: Safety Requirements, IEC 61400-1, IEC Standard, 1994.

[52] R. Yokoyama, S. H. Bae, T. Morita, and H. Sasaki, Multi-objective generation dispatch based on probability security criteria, IEEE Trans. Power Syst., vol. 3, no. 1, pp. 317–324, Feb. 1988.

[53] I. Troen and E. L. Petersen, European Wind Atlas, Riso Nat. Lab., 1989.

[54] L. L. Freris, Wind Energy Conversion Systems. Englewood Cliffs, NJ: Prentice Hall, 1990.

[55] M. Evans, N. Hastings, and B. Peacock, Statistical Distributions, Wiley-Interscience, 2000.

[56] N. L. Johnson, S. Kotz, and N. Balakrishnan, Continuous Univariate Distributions, vol. 2. Wiley-Interscience, 1994.

[57] Patrick F. Dunn. Measurement and data analysis for engineering and science, New York: McGraw–Hill, 2005.

[58] G. W. Corder, D. I. Foreman, Nonparametric statistics for non-statisticians: A step-by-step approach, Wiley, 2009.

[59] N. Ebrahimi, E. S. Soofi and R. Soyer, "On the Sample Information about Parameter and Prediction," Statistical Science, vol. 25, pp. 348-367, 2010.

[60] A. C. Cohen, "Multi-censored sampling in 3 parameter Weibull distribution," Technometrics, vol. 17, no. 3, pp. 347-51, 1975.

[61] J. R. Stedinger, "Fitting log normal-distributions to hydrologic data," Water Resour Res, vol. 16, no. 3, pp. 481-90, 1980.

[62] I. Ntzoufras, Bayesian Modeling Using WinBUGS, Wiley, 2009.

[63] A. H. Shahirinia, E. S. Soofi and D. C. Yu, "Simulation of Economic Dispatch Outputs for Wind-Penetrated Power Systems: a Bayesian Approach", under review at IEEE Trans. Power Systems.

[64] N. Ebrahimi, E. S. Soofi and R. Soyer, "Information Measures in Perspective," International Statistical Review, vol. 78, pp. 383–412, 2010.

[65] M. Asadi, S. Ashrafi, N. Ebrahimi and E. S. Soofi, "Models Based on Partial Information about Survival and Hazard Gradient," Probability in the Engineering and Information Sciences, vol. 24, pp. 561–584, 2010.

[66] J. F. Manwell, J. G. McGowan and A. L. Rogers, Wind energy explained: theory, design and application, 2002.

[67] T. Burton, D. Sharpe, N. Jenkins and E. Bossanyi, Wind energy handbook, Wiley; 2001.

[68] J. E. Oakley and A. O'Hagan, "Probabilistic sensitivity analysis of complex models: a Bayesian approach," J. R. Statist. Soc. B, vol. 66, Part 3, pp. 751–769, 2004.

[69] L. Bayón, J. M. Grau, M. M. Ruiz, and P. M. Suárez, The exact solution of the environmental/economic dispatch problem, IEEE Trans. Power Syst., vol. 27, no. 2, pp. 723-731, May 2012.

[70] A. H. Shahirinia, A. Hajizadeh and D. C. Yu, "Control of a hybrid wind turbine/battery energy storage power generation system considering statistical wind characteristics", Journal of Renewable and Sustainable Energy, vol. 4, issue 5, 2012.

Amir SHAHIRINIA

Summary

. Electrical Engineering PhD student in power systems and renewable energy grid integration. Developed a practical tool to model uncertain renewable-penetrated power systems for renewable energy lab at University of Wisconsin-Milwaukee. [Research Assistant]

. Electrical Engineering PhD student in power systems and renewable energy grid integration. Developed Stochastic Economic Dispatching algorithm for renewable energy lab at University of Wisconsin-Milwaukee. [Research Assistant]

. Engineering intern in power electronics and dedicated team player with extensive knowledge of electrical engineering concepts and a creative aptitude for new electric drive development.

. Electrical Engineering MSc student in power systems and renewable energy grid integration. Developed algorithm for optimal sizing and planning of renewable-penetrated power systems as well for renewable energy lab at K.N.Toosi Univ. of Tech. [Research Assistant]

Highlights

. Power and renewable energy systems

. Power electronics and control

. Data science and statistical analysis

. Hybrid systems

Accomplishments

. Leading member of the new algorithm development team that designed several applications for renewable and sustainable energy lab at UWM.

. Led the Electrical Engineering and Design team in optimal operation and planning of high renewable-penetrated power systems projects that spanned 2005.

Experience

Engineering intern (Rockwell Automation) — **2014-2015**, **Mequon, WI**
. Investigating the application of silicon-carbide (SiC) MOSFET power device to switched-mode power supplies (SMPS) for AC drives.

Research assistant (University of Wisconsin-Milwaukee) — **2010-2014**, **Milwaukee, WI**
. Implemented practical tool to model renewable-penetrated power systems statistically as part of a NSF-funded project.

Education

Ph.D.: Electrical Engineering (University of Wisconsin-Milwaukee)

2014
Milwaukee, WI, USA

Coursework in Advance power electronics II, Wind energy, Control of distributed generations, Control of renewable energy, Bayesian statistics,

Achievements:

. GPA 4:00

. Peer reviewed publication, more than 10 publications

. Chancellor award, Academic Excellence, University of Wisconsin-Milwaukee, Spring 2010, Fall 2010, Spring 2011, Fall 2011, Spring 2012, Fall 2012, Spring 2013, Fall 2013, Spring 2014, Fall 2014

. Outstanding Teaching Assistant of University of Wisconsin Milwaukee, Spring 2011

. Paper Award (2nd rank) ,7th Annual Green Energy Summit, Milwaukee, WI, USA March 24-26, 2010

. Student member of IEEE

M.Sc.: Electrical Engineering (K.N.Toosi University of Technology)

2005
Tehran, Iran

Coursework in Advanced power electronics I, Electric machinery design, Modern control, Inactive and reactive power control, Renewable energy, Advance power systems dynamic, Dispatching

Achievements:

. Peer reviewed publication, more than 10 publications

. Outstanding Research Assistant, Fall 2005

. Outstanding Research Assistant, Fall 2004

. Student member of IEEE

B.Sc.: Electrical Engineering (K.N.Toosi University of Technology)

2003
Tehran, Iran

Achievements:

. Peer reviewed publication, 2 publications

. Student fellowship, Fall/Spring 2000, Fall/Spring 2002

. Best undergraduate thesis award

. Member of IEEE

Publications

Book chapter

[1] A.Hajizadeh, ***A.H.Shahirinia,*** D.C.Yu, Chapter#13 "Power Control of Plug-in Electric Vehicles in Smart Grids", Autonomous Hybrid Vehicles: Intelligent Transport Systems and Automotive Technologies, NOVA Science Publishers, INC. (May. 2014, ISBN: 978-606-560-327-1)

Published journal articles

[2] A.Hajizadeh, ***A.H.Shahirinia***, D.C.Yu, "Self-Tuning Indirect Adaptive Control of Non-inverting Buck-Boost Converter", Journal of Power Electronics (IET), May. 2015 (In press)

[3] A.Hajizadeh, ***A.H.Shahirinia***, D.C.Yu, "Fuzzy Control of Hybrid Diesel Generator/ Fuel Cell/ Energy Storage Power Sources for Marine Power System", Journal of Fuel Cell Science and Technology (ASME), Apr. 2015 (In press)

[4] ***A.H.Shahirinia,*** A.Hajizadeh, D.C.Yu, A.Feliachi, "Control of a Hybrid Wind Turbine/Battery Energy Storage Power Generation System Considering Statistical Wind Characteristics", Journal of Renewable and Sustainable Energy (JRSE), Vol. 4, 2012

[5] P.Naderi, ***A.H.Shahirinia***, O.P.Malik, Power System Stabilization Using Optimal Placement of Stabilizers and Design of Local Robust Controllers," International Review of Automatic Control (IREACO)", Vol. 2. n. 2, pp. 163-169- March 2009

[6] ***A.H.Shahirinia***, A.Radan, "Novel Carrier-Based PWM Methods for Multi-Level Inverters", European Power Electronics and Drives (EPE), Vol. 18, No. 2, 2008

[7] ***A.H.Shahirinia***, A.Hajizadeh, A.R.Moghaddmjoo, "Genetic-Based Size Optimization of Renewable Energy Sources", International Journal of Power and Energy Systems (ACTA Press), Vol. 28, No. 1, 2008

[8] ***A.H.Shahirinia***, S.M.M.Tafreshi, A.Hajizadeh, A.R.Moghaddmjoo, "Optimal Design of Multi-Sources Stand-Alone Hybrid Power System Using Genetic Algorithm", Journal of Iranian Association of Electrical and Electronics Engineers (IAEEE), Vol.3, No.2-Fall and winter 2006

Under review journal articles

[9] ***A.H.Shahirinia***, E. Soofi, D.C.Yu, "Simulation of Economic Dispatch outputs for Wind-Penetrated Power System", International Journal of Electrical Power & Energy Systems (ELSEVIER)

[10] ***A.H.Shahirinia***, R. Tallam, "Simulation Model for Multi-Winding Transformers for Switched Mode Power Supply", IEEE Trans. On Industry Applications

In preparation journal articles

[11] ***A.H.Shahirinia***, E.Soofi, D.C.Yu, "New Probability Models for Wind-Penetrated Power Systems; Bayesian Approach", In preparation for IEEE Transaction Power System

[12] ***A.H.Shahirinia***, E.Soofi, D.C.Yu, "A New Approach in Simulation of Dependent Wind Speeds; Copula Application", In preparation for IEEE Transaction Power System

[13] ***A.H.Shahirinia***, A.Hajizadeh, D.C.Yu, "Optimal Nero-Fuzzy Control for Power Management and Voltage Ride-through Improvement of Hybrid Wind/Energy Storage Power Generation", In preparation for IEEE Transaction on Sustainable Energy

Published conference articles

[14] ***A.H.Shahirinia***, A.Hajizadeh, S.Arabameri, D.C.Yu, "State of Charge Estimation of Battery Energy Storage for Solar Power Systems ", Submitted to 3rd International Conference on Renewable Energy Research and Applications (ICRERA), Milwaukee, USA, 19-22 Oct. 2014

[15] ***A.H.Shahirinia***, A.Hajizadeh, D.C.Yu, "Robust Control of Hybrid Wind / Energy Storage Power Generation System Considering Statistical Wind Characteristics", IEEE International Power and Energy Conference (PECON), Kota Kinabalu, Sabah, Malaysia, 2-6 Dec. 2012

[16] A.Hajizadeh, ***A.H.Shahirinia***, D.C.Yu, "Power Control of Autonomous Hybrid Diesel Generator/ Fuel Cell Marine Power System Combined with Energy Storage", IEEE International Power and Energy Conference (PECON), Kota Kinabalu, Sabah, Malaysia, 2-6 Dec. 2012

[17] ***A.H.Shahirinia***, S.M.M.Tafreshi, A.Hajizadeh, A.R.Moghaddmjoo, "Optimal Sizing of Hybrid Power System Using Genetic Algorithm", IEEE International Conference of Future Power Systems, Amsterdam, Netherlands, Nov. 16-18, 2005

[18] A.Radan, ***A.H.Shahirinia***, M.Falahi, "Evaluation of carrier-Based PWM Methods for Multi-level Inverters", IEEE International Symposium on Industrial Electronics (ISIE 2007), Vigo, Spain, 4-7 Jun. 2007

[19] ***A.H.Shahirinia***, A.Hajizadeh, P.Naderi, A.R.Moghaddmjoo, "The Best Size Planning of a PV/Wind, Local Remote Hybrid Power System", International Conference on Electrical Engineering (ICEE 2006), Tehran, Iran, 16-18 May 2006

[20] P.Naderi, ***A.H.Shahirinia***, S.M.T.Bathaee, Batul Labibi "A New Approach in Decentralized Control of Multi-Machines Large Scale Power System", International Conference on Electrical Engineering (ICEE 2006), Tehran, Iran, 16-18 May 2006

Honors and awards

. **Chancellor award,** Academic Excellence, University of Wisconsin-Milwaukee, Spring 2010, Fall 2010, Spring 2011, Fall 2011, Spring 2012, Fall 2012, Spring 2013, Fall 2013, Spring 2014, Fall 2014

. Outstanding Teaching Assistant, University of Wisconsin Milwaukee, Spring 2011

. Paper Award (2nd rank), 7th Annual Green Energy Summit, Milwaukee, WI, USA March 24-26, 2010

. Outstanding Research Assistant, K. N. Toosi Univ. of Tech., Fall 2005

. Outstanding Research Assistant, K. N. Toosi Univ. of Tech., Fall 2004

Computer skills

. Scientific applications**,** MATLAB/Simulink, Ansys, Simplorer, PSpice, LTSpice, "R"/R Studio, Minitab, WinBUGS, SAS, C/C++

. Familiar packages**,** EMTP, DigSi, HOMER, HYBRID 2, VIPOR, ETAP